Modelling of Personalised Landmarks

Inaugural-Dissertation
for the degree of
Doctor of Natural Sciences (Dr. rer. nat.)
in the
Faculty of Applied Computer Science
University of Augsburg

by
Eva Nuhn

2019

Evaluation committee:	Prof. Dr. Sabine Timpf
	Prof. Dr. Kai-Florian Richter
	Prof. Dr. Jukka M. Krisp
Date of Defence:	October 28th, 2019

Abstract

Numerous studies claim that personal dimensions - such as *personal interests* or *prior spatial knowledge* - influence the identification of landmarks for pedestrians. We assume that the collection of personal data is the highest effort for the identification of personalised landmarks. Therefore, we need to make sure that the data collection effort is justified relative to the benefits that can be accrued through the provision of personalised landmarks.

In this thesis we determine which personal dimensions play a role for the identification of personalised landmarks and focus on *prior spatial knowledge* and *personal interests*. On the basis of these dimensions we personalise existing mathematical models. From the amount of possible models available, we limit our investigations on four of them: the weighted sum model, the weighted product model, a decision flow chart, and a decision tree model. We train and implement the personalised models and use them to identify landmarks selected by participants of a survey. The results of the models are evaluated and compared with statistical methods. In addition, we train conventional, non-personalised models and use them also to identify the landmarks selected by the participants. We compare the results of both models, personalised and conventional, to see if there are advantages of personalisation.

The comparison shows that although the personalised models respond sensitively to personal dimensions, a personalised model does not identify significantly more landmarks selected by survey participants than a conventional model. This means that the collecting of personal data is unlikely to justify the effort. Therefore, it is most likely sufficient to focus on existing conventional, non-personalised models and to concentrate on their use in applied pedestrian wayfinding applications.

Landmark Identification, Personalisation, Spatial Knowledge, Personal Interests

Zusammenfassung

Zahlreiche Studien behaupten, dass persönliche Dimensionen - wie *persönliches Interesse* oder *vorheriges räumliches Wissen* - die Identifikation von Landmarken für Fußgänger beeinflussen. Wir gehen davon aus, dass der größte Aufwand zur Identifikation personalisierter Landmarken bei der Erhebung personenbezogener Daten entsteht. Deshalb muss sichergestellt werden, dass der Aufwand für die Datenerhebung im Vergleich zu den Vorteilen, welche die Bereitstellung personalisierter Landmarken mit sich bringt, gerechtfertigt ist.

In dieser Arbeit ermitteln wir, welche persönlichen Dimensionen eine Rolle bei der Identifikation personalisierter Landmarken spielen und konzentrieren uns auf *vorheriges räumliches Wissen* und *persönliche Interessen.* Auf Basis der Dimensionen personalisieren wir bestehende mathematische Modelle. Aus der Fülle möglicher Modelle beschränken wir unsere Untersuchungen auf vier ausgewählte: das gewichtete Summenmodell, das gewichtete Produktmodell, ein Entscheidungsflussdiagramm und ein Entscheidungsbaummodell. Wir trainieren und implementieren die personalisierten Modelle und verwenden sie, um die von den Teilnehmern einer Studie ausgewählten Landmarken zu identifizieren. Die Ergebnisse der Modelle werden mit statistischen Methoden ausgewertet und verglichen. Des Weiteren trainieren wir konventionelle, nicht personalisierte Modelle und verwenden sie ebenfalls zur Identifikation der von Teilnehmern ausgewählten Landmarken. Wir vergleichen die Ergebnisse beider Modelle, personalisiert und konventionell, um festzustellen, ob die Personalisierung Vorteile bietet.

Der Vergleich zeigt, dass obwohl die personalisierten Modelle sensitiv auf persönliche Dimensionen reagieren, ein personalisiertes Modell nicht signifikant mehr von den Studienteilnehmern ausgewählte Landmarken als ein konventionelles Modell identifiziert. Dies bedeutet, dass der Aufwand für die Erhebung personalisierter Daten vermutlich nicht gerechtfertigt ist. Daher ist es höchstwahrscheinlich ausreichend, sich auf vorhandene konventionelle, nicht personalisierte Modelle und auf deren Verwendung in Wegfindungsanwendungen für Fußgänger zu konzentrieren.

Landmarkenidentifikation, Personalisierung, Räumliches Wissen, Persönliche Interessen

Acknowledgments

I would like to acknowledge the help and support of many people. First of all, I am grateful to Prof. Dr. Sabine Timpf for giving me the opportunity to come back to University, for encouraging me, and for helping me to develop what has become this thesis. In particular I want to thank her because she was always ready to discuss problems with me and to find solutions. I am grateful to Prof. Dr. Kai-Florian Richter for reviewing my thesis and for giving me valuable comments at conferences and workshops. I thank Prof. Dr. Jukka Krisp for reviewing my thesis and for giving me valuable suggestions during numerous GIG-meetings. I also thank Prof. Dr. Ulrike Ohl for examining my thesis.

I am thankful to all current and former colleagues in our Geoinformatics Group for their support: Maja, with whom I shared my office and who was always ready for discussion and advice. Christian, who supported me with the data acquisition for the survey. Linfang and Jean, with whom I shared the VisLab when I just started at the University in Augsburg. Andreas and Johannes, who were always open for questions and discussions.

I would like to thank the two students of my project seminar who helped to collect the data for the survey, to acquire participants, and to interview them. And, of course, I thank all the participants who were willing to complete the survey.

I am grateful to my former math teacher, who accepted me at high school, and who unfortunately passed away one month after my defence. Although he never knew, I wouldn't be here today without him.

A very special gratitude goes out to my former colleagues from the UniBW - the place where the idea of writing a PhD thesis was born. Especially I want to thank Manu, Admire, Agnes, Stephan, Iris, and Pietzi and, of course, Prof. Reinhardt.

I thank my family and friends. My parents for always supporting and encouraging me to go my own way. My mother who helped to finish this thesis by improving the English wordings. My parents in law who were there when I was away for conferences. My friends for the great time we have together.

My biggest thanks go to my two wonderful girls. They developed very creative ways to keep themselves entertained while I was working on this thesis instead of spending time with them. Last but not least I want to thank my husband Matthias who never lost patience with me. This made it possible for me to meet deadlines and finally to finish this thesis that otherwise would have been impossible.

In a system that automatically generates route directions, not including landmarks is a violation of cognitive ergonomics.

Richter & Klippel (2007, p. 459)

Contents

List of Figures

List of Tables

Abbreviations

avgSI average sensitivity index

CART Classification and Regression Trees

CdFc Conventional Decision Flow Chart

CdTm Conventional Decision Tree Model

cult cultural

CwPm Conventional Weighted Product Model

CwSm Conventional Weighted Sum Model

dFc Decision Flow Chart

DP decision point

dTm Decision Tree Model

gast gastronomy

hist historical

iLM landmark interest

LM landmark

NAL not a landmark

OSM OpenStreetMap

PdFc Personalised Decision Flow Chart

PdTm Personalised Decision Tree Model

pInt personal interests

POI Point of Interest

PspK prior spatial knowledge

PwPm Personalised Weighted Product Model

PwSm Personalised Weighted Sum Model

sem semantic

shop shopping

SI sensitivity index

str structural

VGI Volunteered Geographic Information

vis visual

wPm Weighted Product Model

wSm Weighted Sum Model

Chapter 1

Introduction

> *'Go straight ahead to the small square where the kiosk is, then turn into Karl-Strasse and go past the old garden, where once the mining director lived. Then - just behind the allotment gardens you will find the scrap yard'(translated from Vahle (2014)).*

This is what Anne's grandmother in the famous children's book *Anne Kaffekanne* (Vahle 2014) says when Anne wants to get to the scrap dealer where she hopes to find her missing coffee pot. Here Anne's grandmother gives route directions which are enriched with *personalised landmarks* - salient objects that are suitable to Anne's needs. The grandmother knows that Anne is familiar with the area and knows where the mining director used to live. Furthermore, the grandmother is aware of the fact that Anne knows the kiosk where she sometimes buys something there for herself. In fact, Anne finds the way without any problems and without getting lost. Even though 'there was little to see of the old garden and the house' (translated from Vahle (2014)) of the mining director.

Anne's grandmother has intuitively chosen route directions with personalised landmarks perfectly fitting her granddaughter. They are tailored to her *prior spatial knowledge* and adapted to her specific and *personal interests*. Today's pedestrian wayfinding applications are far from providing such personalised landmarks. Most of them use the same strategies as car navigation systems and include only a pedestrian mode. They provide route directions tailored to the needs of car drivers but not adequate for pedestrians. The resulting route directions consist of compass directions, distances, and street names. Such a wayfinding system would give Anne the following route direction:

> *'Head north on the square and then continue onto Karl-Strasse. After 177 meters the destination will be on the right.'*

We think most of You can agree that with these route directions Anne would not have found the way as easily as with those from her grandmother. Today's pedestrian wayfinding applications rarely include landmarks. Nevertheless, we know that humans prefer them for navigation (Chapter 2). There are even models available which identify objects that are suitable as a landmark, i.e. objects that 'stick out of their surrounds and, thus, may be assigned a *landmarkness* property' (Richter 2017, p. 136). A personalised landmark may have different aspects - so-called *dimensions* - to explain its *landmarkness*. Currently available *conventional landmark identification models* consider only *landmark dimensions*, which are static and dependent on an object itself. They identify landmarks that are used for route directions of the form:

> *'Go straight ahead to the small square with the monument, then follow the Karl-Strasse until the house with the stucco façade. Then - just behind the grey house you will find the scrap yard'.*

A route direction that most likely makes Anne neither curious nor can guarantee that she knows what a *stucco façade* is. These landmark identification models obviously miss *personal dimensions* which depend on Anne's *prior spatial knowledge* and *personal interests*.

However, what grandmother has done so easily is rather hard for an automated system to replicate. To be able to generate route directions with personalised landmarks, an application would have to collect a lot of personal data about Anne. It has to know where Anne has been before and that she likes to buy something at the kiosk. In order to find that out the application would have to ask Anne a number of questions before she would be able to receive route directions. Anne would, very likely, be unwilling to answer all these questions, because she wants to reach the scrap yard as soon as possible.

The highest cost for the provision of personalised landmarks is personal data collection. Therefore, we need to make sure that the data collection effort is justified in relation to the advantages that can be achieved through the provision of personalised landmarks. Currently, neither we know which personal dimensions play a role for the identification of personalised landmarks nor whether these dimensions have an impact on personalised landmark identification. Therefore, our aim is to find out whether a *personalised landmark identification model* that incorporates *prior spatial knowledge* and *personal interests* identifies more landmarks selected by humans than a conventional, non-personalised model. If this is not the case, it is most likely sufficient to focus on existing conventional, non-personalised models and to concentrate on their use in applied pedestrian wayfinding applications.

1.1 Hypothesis and Goals

We define landmarks as salient objects that may attract our attention (Richter & Winter 2014) (for a detailed definition of landmarks see Section 2.1.2). The property that turns a conventional geographic object into a landmark is called *salience* (Raubal & Winter 2002, Elias 2003*b*). There is a general distinction between approaches to landmark identification and landmark integration (for details see Section 2.2). Landmark identification models concern the assessment of object salience for navigation and result in a pool of potential landmarks (Richter & Winter 2014). Currently available *conventional landmark identification models* consider only *landmark dimensions*, which are static and dependent on an object itself. We make a contribution to landmark identification and develop *personalised landmark identification models* to support the creation of *cognitively ergonomic route directions* (Klippel et al. 2009). Such route directions consider aspects of personalisation such as the 'user's familiarity with an environment, as well as personal styles' (Klippel et al. 2009, p. 231). Therefore, we focus on *prior spatial knowledge* of a traveller and *personal interests* as important *personal dimensions* in our *personalised landmark identification models.* Our hypothesis is the following:

> *A personalised landmark identification model that incorporates prior spatial knowledge and personal interests identifies more landmarks selected by humans than a conventional, non-personalised model.*

In order to test this hypothesis we investigate possible personal dimensions, formalise and implement personalised landmark identification models, and compare their results to conventional models. The following steps help to reach these goals:

1. Investigate dimensions playing a role for personalised landmark identification with a focus on *personal interests* and *prior spatial knowledge.*

2. Investigate salience measures of the personal dimensions, *personal interests* and *prior spatial knowledge.*

3. Develop landmark identification models both conventional and personalised.

4. Implement the landmark identification models.

5. Collect data for landmark dimensions and collect personal dimensions and landmarks in the framework of a survey to feed the landmark identification models.

6. Create the landmark identification models with the help of the collected data.

7. Test the performance of the models on collected data and compare the models results (identified landmarks) with landmarks selected by the participants of the survey.

8. Perform a sensitivity analysis to identify the dimensions which influence the results of the personalised landmark identification models.

9. Compare the results of the personalised landmark identification models with the results of the conventional landmark identification models.

1.2 Approach

A personalised landmark may have different dimensions to explain its landmarkness. The first step of our approach is the investigation of these dimensions playing a role for personalised landmark identification. Winter et al. (2012) identify the need for context-dependent identification of landmarks focusing on:

1. the context that represents the appearance, the efficacy from all directions, or cultural importance of the landmark itself and

2. the context that represents preferences of the traveller as an individual (e.g. mobility, gender, age, education, or home town).

The former is static and dependent on an object itself. The latter changes with each individual traveller, because whether an object becomes a landmark is not only affected by the object itself but also by the perspective of the traveller (Caduff & Timpf 2008). We differentiate the following dimensions based on Winter et al.'s (2012) work:

1. landmark dimensions and

2. personal dimensions.

We build on the definitions of Sorrows & Hirtle (1999) and Raubal & Winter (2002) for landmark dimensions. These are the visual, the semantic, and the structural dimension. Additionally, we add a *landmark interest* dimension to consider the topic of interest. We identify attributes and attribute values for the landmark dimensions.

There are several personal dimensions: *personal knowledge*, *personal interests*, *personal goals*, *personal background*, and *individual traits* (Brusilovsky & Millán 2007). We focus on *prior spatial knowledge* and *personal interest* in this thesis. We investigate attributes and attribute values for these personal dimensions.

Furthermore, we investigate salience measures of the dimensions. We adapt existing salience measures from Raubal & Winter (2002) and Nuhn et al. (2012) for the landmark dimensions. We introduce a salience measure for the *landmark interest* dimension. In addition, we investigate methods to calculate salience of the personal dimensions and develop salience measures for *prior spatial knowledge* and *personal interests*.

Then, we develop landmark identification models. The conventional models built on landmark dimensions and the personalised models add the personal dimension. Due to the amount of possible models which could be used as basis we limit our investigations to three mathematical models that are based on theory and a machine learning model that has the ability to learn from data (Samuel 1959):

1. Models based on theory

    - **Weighted Sum Model (wSm)** Raubal & Winter (2002) propose a wSm for modelling landmark salience. This model is widely used for landmark identification (e.g. Winter (2003), Nothegger et al. (2004)). We use this model as a Conventional Weighted Sum Model (CwSm) and extend this model with personal dimensions (Personalised Weighted Sum Model (PwSm)). The result of these models is an overall measure of landmark salience for an object.
    - **Weighted Product Model (wPm)** A wPm is an alternative to a wSm. We are not aware of an existing wPm for landmark identification. We build a Conventional Weighted Product Model (CwPm) and a personalised model (Personalised Weighted Product Model (PwPm)). The result of the PwPm and the CwPm is an overall landmark salience measure for an object.
    - **Decision Flow Chart (dFc)** There is a long tradition of using diagrams for a large variety of tasks. Such flowcharts are based on knowledge of

experts or literary research and involve decisions and processes. Since a dFc for landmark identification does not exist, we build both a Conventional Decision Flow Chart (CdFc) and a Personalised Decision Flow Chart (PdFc). The models result in one or more landmarks for a decision point.

2. Machine Learning Model

   - **Decision Tree Model (dTm)** There are numerous machine learning methods. One model is the dTm, which is similar to a dFc but it does not consider processes and concentrates only on decisions and their results. dTms are already used for landmark identification (Elias 2006). We create a Conventional Decision Tree Model (CdTm) and a personalised one (Personalised Decision Tree Model (PdTm)). The resulting models are able to classify objects in *landmark (LM)* and objects which are *not a landmark (NAL)*.

This results in eight models: four conventional and four personalised landmark identification models.

The next step of our approach is the implementation of these models. We implement all the models and methods using ESRI's ArcGIS 10.5.1 together with Python toolboxes using Python 2.7.13. The python site package ArcPy provides an environment for developing Python scripts and enables writing customised ArcGIS applications and scripts (ESRI 2018). In addition, we use several packages e.g. the statistical packages *scipy.stats* and *scikit-learn* which provide simple and efficient tools for data mining and data analysis (Pedregosa et al. 2011).

We collect data for landmark and personal dimensions. While landmark dimensions are extracted from official databases or acquired during field surveys, personal dimensions are collected by a survey. We perform the survey in the inner city of Augsburg and ask participants to select landmarks (LM) and objects which are not landmarks (NAL). Furthermore, we ask them to provide information on personal dimensions.

We use a part of the collected data to create (*train*) the landmark identification models (*training set*) and the other part to test the models' performances (*test set*). The machine learning models, CdTm and PdTm, learn their model parameters from the *training set*. The conventional models based on theory have no unknown model parameters, whereas the model parameters of the personalised models that are also based on theory need to be identified. This concerns the weights of both: PwSm and the PwSm and the flow of the PdFc as well.

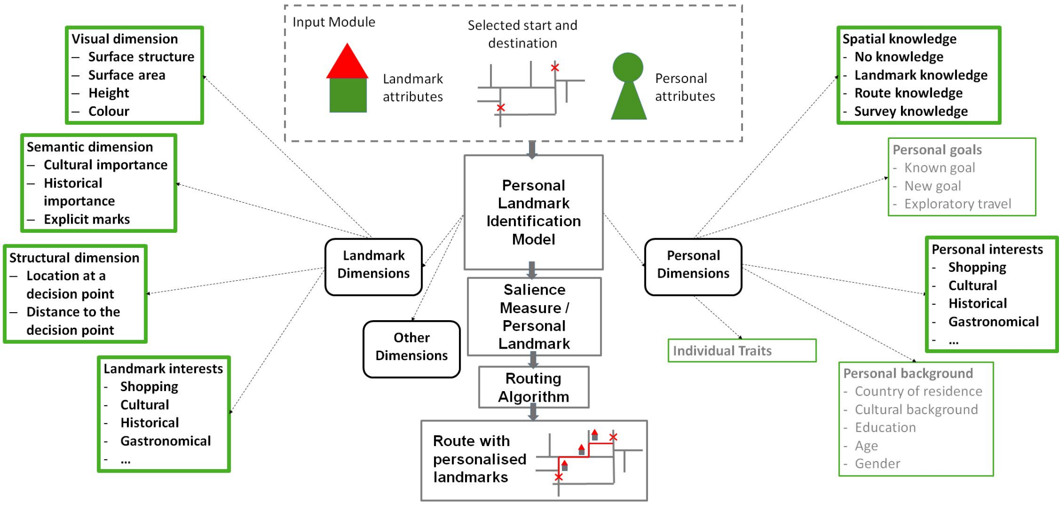


Figure 1.1: Architecture of a pedestrian wayfinding application.

After the models are created we have conventional and personalised models that are able to identify whether a sample of a new unseen dataset is a LM or a NAL (*testing*). We identify how the models perform on the test set and compare their identified landmarks with landmarks selected by the participants of the survey. We compare the results of the models and identify which of the models deliver better results.

Furthermore, we perform a sensitivity analysis to identify the dimensions which influence the output of the personalised models. We vary one dimension at a time to investigate the effect that changes in dimensions have on the outputs of the models.

The last step of our approach focuses on the comparison of the results of the personalised landmark identification models with the results of the conventional landmark identification models. We perform an analysis to find out whether there are statistically significant differences between the conventional and the personalised models.

1.3 Scope and Methods

The general idea of this thesis is to integrate the personalised landmark identification models in a pedestrian wayfinding application generating routes with personalised landmarks. Figure 1.1 shows a possible architecture of such an application with an input module, the personalised landmark identification model, as well as the processing of the result of the model in a routing algorithm. The input module of such an application would need three different input types to output a route with personalised landmarks:

1. values of the landmark dimensions,
2. start and destination of the route, and
3. values of the personal dimension.

The personalised landmark identification models allow the determination of an overall landmark salience measure for objects or a classification of objects in landmarks or objects that are not landmarks. This result will then be integrated in the generation of a route between the defined start and destination, i.e. it can be introduced in a shortest path algorithm. The result of the routing algorithm is an optimal route in terms of personalised landmarks.

Within the scope of this thesis we concentrate on personalised landmark identification models for pedestrians in European urban inner city environments. We do not include other environments or other modes of transport in this thesis.

In this thesis a landmark might be any urban structure (buildings as well as e.g. monuments). That means, this work is not restricted to buildings but also treats other geographic objects in an urban environment (e.g. water wheels, information panels, or dust bins). There is a distinction between local and global landmarks in landmark research (for details see Section 2.1.2). In this thesis we focus exclusively on local landmarks. Another differentiation is in two and three-dimensional landmarks (for details see Section 2.1.2). Here, the focus is on local three-dimensional landmarks. This might be either a building such as a shop, a restaurant, or a school but also towers, city gates, or city walls or even point-like objects, such as street signs, bus stops, or advertisement pillars.

We assume that a street network consists of nodes (street intersections) and edges (street segments). A decision point (DP) is a node of a street network where actions (e.g. re-orientations) are performed (Klippel & Winter 2005). We consider objects at decision points and assume that there is at least one object at each decision point available as input for our models. We do not consider objects along street segments in this thesis, although the landmark identification models might be transferred. We divide objects in *landmark (LM)* and *not a landmark (NAL)* at decision points.

We collect personal information in the framework of a survey. We control that the participants are not visually impaired or disabled because then their information needs and their landmark selection would differ (Golledge et al. 2000, Loomis et al. 2001).

We use mathematical models as basis for the conventional landmark identification models. There might be methods to improve the results of these conventional models. However, the improvement of the conventional models is not within the scope of this thesis.

1.4 Expected Result and Contributions

The main contribution of this thesis is an answer to the question whether a personalised landmark identification model incorporating *prior spatial knowledge* and *personal interests* identifies more landmarks selected by humans than a conventional, non-personalised landmark identification model. In order to achieve this result this thesis makes the following contributions:

- An analysis of personal dimensions playing a role for personalised landmark identification. We contribute attributes and attribute values for the identified dimensions.
- An analysis of salience measures for *prior spatial knowledge* and *personal interests* forming the basis for personalised landmark identification models.
- An adaptation of mathematical models for landmark identification - conventional and personalised - in the context of pedestrian wayfinding applications.
- An implementation of the conventional and personalised landmark identification models.
- A survey for the collection of personalised landmarks and personal dimensions.
- A comparison of the different models and their ability to identify landmarks selected by participants of a survey.
- A sensitivity analysis to identify attributes of landmark and personal dimensions influencing the output of the personalised landmark identification models.
- A statistical evaluation and comparison of the results of the personalised landmark identification models in contrast to the results of conventional landmark identification models.

1.5 Thesis Outline

The structure of this thesis is as follows. In the next chapter (Chapter 2) we discuss the theoretical background and related work. We review what has already been done in the field of landmarks and human wayfinding and give a detailed definition and characterisation of landmarks (Section 2.1). Furthermore, we look at existing methods for modelling landmarks for route directions (Section 2.2) and review research towards personalised landmarks (Section 2.3). Chapter 2 concludes with implications for the modelling of personalised landmarks (Section 2.4).

Chapter 3 introduces mathematical models as well as analysis methods. We intend to use models that are based on theory and a machine learning model for landmark identification (Section 3.1). In Section 3.2.1 we discuss the traditional approach to train and test the machine learning models. Inspired by this approach, we investigate in Section 3.2.2 methods to 'train' and test the models based on theory. Furthermore, we elaborate on how to perform sensitivity analysis of the models in order to evaluate the effects of the inputs on the models' behaviour (Section 3.3). Afterwards, we evaluate a comparison method for the results of the models (Section 3.4). We close with an outlook on the study setup of this thesis (Section 3.5).

Chapter 4 deals with landmark identification models. It investigates which dimensions play a role for the inclusion in such models and identifies attributes and possible attribute values for landmark and personal dimensions (Section 4.1). We investigate salience measures for the attributes of these dimensions (Section 4.2). Finally, we describe in Section 4.3 the conventional and the personalised models to calculate the overall salience of an object.

The focus of Chapter 5 is on data collection and preparation. We describe how we extract attributes for the landmark dimensions from official databases or how we acquire them via field surveys (Section 5.1.1). Furthermore, we describe the setting of a survey in the inner city of Augsburg to collect personal information and landmarks (Section 5.1.2). Then, we calculate salience measures for the collected data and present the input data for the models (Section 5.2). We divide our collected dataset into a training set and a test set (Section 5.3). We conclude the chapter with the calculation of overall landmark salience according to the conventional models and compare the results (Section 5.4).

In Chapter 6 we describe in detail how we carry out the training and testing of the personalised landmark identification models. First, we describe how to train the models with the collected dataset from Chapter 5 (Section 6.1). We use the traditional approach for the machine learning model and the approach based on the traditional approach for the other models respectively. Then, we use the models to identify landmarks and present the results (Section 6.2). We close the chapter with a discussion of the results of the training and the testing of the personalised models (Section 6.3).

We start Chapter 7 with a sensitivity analysis of the personalised models (Section 7.1). Furthermore, we carry out a statistical evaluation of the model results and compare them to the results of the conventional models (Section 7.2). Finally, we draw conclusions of the results of the sensitivity analysis and the comparison of the models with regard to our hypothesis (Section 7.3).

In Chapter 8 we discuss the results and investigate a number of reasons for them: not considered further dimensions (Section 8.1), the methods to calculate salience measures (Section 8.2), the models to calculate overall salience (Section 8.3), the dataset (Section 8.4), and the survey design (Section 8.5).

Chapter 9 summarises the research of this thesis (Section 9.1). We present results and conclusions (Section 9.2) and present ideas for future research on the modelling of personalised landmarks (Section 9.3). Finally, we conclude with some remarks (Section 9.4).

The appendix shows tables and figures that do not add to the argument being made in the main text but should be included for the sake of completeness (Appendix A).

Chapter 2

Related Work

This chapter presents the previous work in pedestrian wayfinding and landmark research. The first Section 2.1 explores theories about landmarks in human wayfinding. The second Section 2.2 investigates methods to identify landmarks and methods to integrate them in route directions. The third section 2.3 reviews research towards personalised landmarks. The chapter concludes with a summary and implications for the modelling of personalised landmarks (Section 2.4).

2.1 Landmarks in Human Wayfinding

This section highlights cognitive aspects of human wayfinding and then evaluates definitions and characteristics of landmarks.

2.1.1 Cognitive Aspects of Human Wayfinding

In this thesis we address a topic part of the research work in the field of location based services and pedestrian wayfinding applications (Huang et al. 2018). Human wayfinding research is part of cognitive science. It investigates the process that takes place when humans orient themselves and navigate through the environment (Raubal & Winter 2002). Various theories try to explain how people find routes in space, which information they need to find these routes and how they communicate route directions (Allen 1997, 1999, Daniel & Denis 1998, Golledge 1999, Kuipers 1978). The following sections elaborate on these core points of human wayfinding.

Wayfinding Definition, Tasks, and Means

'Wayfinding describes a person's ability, both cognitive and behavioral, to reach spatial destinations' (Passini 1984, p. 154). According to Montello (2005) *human* *wayfinding* is one of two components of navigation. He defines it as a 'goal-directed and planned movement of one's body around an environment' (Montello 2005, p. 259). It requires from people to solve problems and to make decisions (e.g. choosing routes, creating shortcuts, scheduling trips) (Montello & Sas 2006). In contrast, *locomotion*, the second component, requires coordination with the local or near surroundings directly accessible to sensory and motor systems of humans (Montello 2005, Montello & Sas 2006).

People pursue different goals while travelling through space. Wiener et al. (2009) subdivide unaided wayfinding into travelling with a specific spatial goal and travelling with a non-spatial goal. Allen (1999) categorises wayfinding tasks according to functional goals:

- **Travel with the goal of reaching a familiar destination** The focus of travelling with a specific spatial goal is primarily on reaching a particular destination. This is a very common task. An example from everyday life is commuting between home and work place (Allen 1999).
- **Travel with the goal of reaching a new destination** In this case the specific spatial goal is unknown. Travelling towards an unfamiliar destination is mostly carried out with different kinds of wayfinding aids (e.g. maps or verbal route directions, for an overview see Elias (2006)).
- **Exploratory travel** In contrast to travelling with a specific spatial goal in mind, the reason for travelling with a non-spatial goal is for example to explore a new environment (e.g. after moving to a new town, for touristic issues, or just to walk on the beach (Wiener et al. 2009)). The goal of exploratory travel is to discover and to return to the starting point (Allen 1999).

These three wayfinding tasks can be accomplished by a variety of means (Allen 1999). One means is piloting between landmarks. This method is equally applicable to all three wayfinding tasks. It is an efficient means for reaching familiar or new destinations. The success of this means is dependent on the recognition of landmarks and remembering the spatial relations between them (Allen 1999). In exploratory wayfinding the traveller selects landmarks rather than relying on familiar or prescribed ones (Allen 1999).

Which Information Do People Need to Find Routes in Geographic Space?

Wayfinding includes determining and following a route between a starting point and a destination (Golledge 1999). In order to find routes in space travellers must have a mental representation of the environment of the route. These mental representations are called *cognitive maps* (Tolman 1948, Downs & Stea 1974, O'keefe & Nadel 1978). Individuals acquire spatial arrangements and navigation possibilities while moving through space and acquire environmental knowledge by experiencing or interacting with the environment (Golledge 1991). Spatial knowledge as a part of a cognitive map is commonly divided into three stages of knowledge with interdependent contents: *landmark*, *route*, and *survey knowledge* (Siegel & White 1975).

- **Landmark knowledge** During spatial knowledge acquisition, landmarks are the first spatial cues that are available in no particular order on a cognitive map (Couclelis et al. 1987). Landmark knowledge is the knowledge where solely landmarks are remembered (Schmauks 1998). This means that only outstanding geographical elements (namely landmarks) in a disordered form are available in memory (Elias 2006).

- **Route knowledge** The knowledge how to get from one place (or one landmark) to another is called route knowledge (Wender 1998). It includes a fixed sequence of locations or landmarks as experienced in traversing a route (Werner et al. 1997). Route knowledge consists of information about the order of landmarks along a route and knowledge of directions such as 'continue straight on' or 'turn right' inbetween those landmarks (Montello 1998).

- **Survey knowledge** Survey knowledge integrates knowledge from different experiences into one single model (Wender 1998). It is the result of the mental integration of two or more routes (Herrmann et al. 1998). This is in contrast to route knowledge which is related to only one route. With the availability of survey knowledge, new routes can be detected and shortcuts can be generated through the environment (Schmauks 1998). Survey knowledge is usually generated from route knowledge through integration into a cognitive map (Herrmann et al. 1998).

The individual stages of spatial knowledge are acquired with different goals, requirements, or tasks influencing the completeness of the spatial knowledge. It can be assumed that spatial knowledge acquisition is faster under time pressure and high attention than if it is acquired implicitly and incidentally (Herrmann et al. 1998). It can be acquired unintentionally in such a way that travellers are not able to indicate how and why they acquired it (Perrig et al. 1993). Spatial knowledge is a dynamic component, thus, it may change over time. Travellers learn new environments thereby increasing their spatial knowledge and getting more familiar with an environment. Likewise, the familiarity may decrease when the traveller has not visited the environment for a longer time and objects did change during this time period (e.g. marks on geographic features, new buildings, disappeared objects, ...). The links between objects and location are lost over time because people forget the environment over time (Pertzov et al. 2012). Spatial knowledge can become elaborate and extensive and supports wayfinding and direction giving (Montello 1998). The three stages framework of Siegel & White (1975) follows the idea that landmark knowledge is a prerequisite for route knowledge, which again is mandatory for survey knowledge (Ishikawa & Montello 2006). Montello (1998) identifies this as a problem of what he called the *dominant* framework from Siegel & White (1975) and offers a more conceptually coherent one. He postulates different types of knowledge that are acquired simultaneously. According to Montello (1998) spatial knowledge is quantitatively accumulated and continuously refined, starting at the point of first exposure to the environment. Because of this idea of continuous acquisition of spatial knowledge in new environments this framework is referred to as the *continuous* framework.

How Do People Communicate Route Directions?

'Route directions are a form of procedural discourse that exploits a vast domain of human knowledge, spatial knowledge, and intends to have other people construct new knowledge to guide their action in the environment' (Denis et al. 1999, p. 171). They are answers to a question of the kind 'How do I get from the university to the station?'. Once persons have acquired spatial knowledge of an environment they are able to give detailed descriptions of a specific route to a traveller. The communication of route directions can be divided in four phases: *initiation*, *route description*, *securing*, and *closure* (Allen 1997). The initiation phase starts with a question e.g. 'How can I get to the train station?' from a traveller to a respondent. Such questions include elements such as the point of origin and constraining conditions. The initiation phase often includes a destination query or a state-of-knowledge query (e.g. 'Ho well do you know the inner city?'). The second phase of route communication is the route description itself, which is then followed by a securing phase, which includes the travellers reaction to the directions (Allen 1997). Clarification queries and confirmation statements are followed by the closure phase, which is a social convention that allows both parts to end the communication (Allen 1997).

Lovelace et al. (1999) identify three major steps as the central part of route directions. The first step is the activation of the spatial knowledge of the environment to be described at an appropriate scale. This knowledge is presumed to be stored in a non-linguistic format. The second step is the choice of a specific route. Lovelace et al. (1999) mention several selection criteria, such as the mode of travel, the desired route characteristics, and the expected spatial knowledge of the traveller (i.e. familiarity with the environment). The last step is the translation of the chosen route into verbal directions (Lovelace et al. 1999) which are communicated to the traveller.

The route directions itself include four characteristics (Daniel & Denis 1998): their *function*, their *content*, their *structure*, and the *perspective* they impose on their users (usually an egocentric perspective). The main function of route directions is to explain to a person the route to a desired goal in a particular environment (Daniel & Denis 1998). They usually contain a number of instructions for behaviour, such as 'turn right', 'going up', 'looking for'. In addition, they include an object or place, which specifies where the behaviour should take place (Daniel & Denis 1998, Passini 1984). Route directions consist of two basic actions: *locomotion* and *reorientation* (Daniel & Denis 1998). Locomotion is needed to reduce the distance between the current position and the destination. Reorientation describes the reduction of the angle between the current direction and the direction to the destination (Daniel & Denis 1998). However, route directions do not only consist of these two basic actions. Consider the following example: 'Proceed 15 meters; stop; rotate 90 degrees to the right; proceed 25 meters; stop; rotate 45 degrees to the left; proceed 20 meters. You are here.' (Daniel & Denis 1998, p. 46). Such directions are very detailed and precise and would be highly useful for guiding a robot. But a person with perceptual access to the environment would never use or produce such directions (Daniel & Denis 1998). Even the shortest direction communicated by a human includes various elements, e.g. landmarks, turns, or descriptive information (Lovelace et al. 1999).

Landmarks in Route Directions

A number of studies deal with route directions and their elements. They show that landmarks are of major importance. The frequency with which they are mentioned is dependent on individual differences. For instance, there are studies clearly indicating that women use landmarks more frequently to describe routes than men do (Dabbs Jr et al. 1998, Galea & Kimura 1993, Sandstrom et al. 1998, Choi et al. 2006, Wang et al. 2019). Independent of these individual differences, landmarks are nearly always used in route directions (Allen 2000, Fontaine & Denis 1999, Michon & Denis 2001). Michon & Denis (2001) even show that directions without landmarks are negatively perceived. In an experiment people were given only minimal information on a route they had to walk. The directions were limited to procedural information, referring to street names and directions (e.g. 'take the street Saint-Antoine on your right'). After walking the route participants were asked to write down their difficulties with these limited descriptions and to suggest possible solutions. Most solutions were related to landmarks and descriptive elements along the route, such as length specifications or the number of roads that had to be passed (Michon & Denis 2001).

Tversky & Lee (1999) show that landmarks are used in verbal route directions as well as in drawn maps. They conducted a survey on an university campus in which passengers were stopped and asked if they knew the route to an off-campus fast food restaurant. If they answered affirmatively, they were asked to give either a short written description of the route or to sketch a map. The results varied, especially the written descriptions. While some of the participants only mentioned essential turning directions, others used complete sentences with detailed landmark descriptions. In fact, more than 90% of the sketch maps and directions included additional information, such as arrows, distances but especially landmarks and landmark descriptions.

Tom & Denis (2003) state that landmarks work better than street signs for wayfinding. They report an experiment where participants were either equipped with street-based directions or with landmark-based directions. Route directions referring to landmarks appeared to be more effective than those referring to streets. A second experiment showed that when people generate route directions they do include more landmark descriptions than references to street names. Finally, Tom & Denis (2003) state that although street names offer an ideal reference in route directions, they appear to be poor guides in contrast to landmarks.

2.1.2 Definition and Characterisation of Landmarks

Many definitions, characterisations, and categorisations of landmarks have been made over the years and a satisfactory one is somewhat elusive (Presson & Montello 1988).

What is a Landmark?

One of the most fundamental definitions of a landmark is introduced by Lynch (1960). In an experiment he asked participants to sketch their home town. Comparing the results, Lynch (1960) identified five basic elements in the *Image of the City*:

1. **Paths** Channels used by pedestrians customarily, occasionally, or potentially. These are streets, side walks, canals, rail roads, and other channels on which people travel. Lynch (1960) shows that paths are the predominant elements

in people's images of the city. All other elements are arranged and related to them.

2. **Edges** Linear elements that are not paths but boundaries. Examples are walls, rail road cuts (which cannot be crossed), or shores. Such boundaries close one area off from the other or define a line along which two regions are related and joined together. They are not as dominant as paths but essential for people to organise objects within a city.

3. **Districts** A two-dimensional medium for a large area in a city, e.g. a part of a city which shares common design elements and identifying characteristics. Individuals are able to enter and leave these areas. The city is structured mostly in districts, individual differences depending on whether paths or districts are the predominant elements.

4. **Nodes** Points and spaces which the traveller can physically enter. Nodes are primarily junctions i.e. a crossing or convergence of paths. The concept of nodes is strongly connected to the concept of paths as junctions define the convergence of paths.

5. **Landmarks** 'Point references considered to be external to the observer, are simple physical elements which may vary widely in scale' (Lynch 1960, p. 78). Landmarks might be buildings, signs, stores, mountains, or other geographic objects.

Years after Lynch (1960), Presson & Montello (1988) state that everything standing out from a scene can be a landmark. Whether an object becomes a landmark is not only affected by the object itself but by the perspective of the observer, the surrounding environment, and the other geographic objects involved (Caduff & Timpf 2008). The number of geographic objects that become a landmark depends as much on how familiar an observer is with the surrounding environment as upon the objects themselves (Lynch 1960). Different observers find different objects to be most useful as a landmark in a given situation (Götze & Boye 2016). Additionally, different people perceive the significance of an object in different ways (Krisp 2016). Most people would agree that the Eiffel Tower is a landmark, however, not so many would agree that the postbox at the street corner is a landmark (Richter & Winter 2014). Thus, the Eiffel tower is a prototype of a landmark (Rosch 1973, Rosch et al. 1976, Rosch 1978), while the postbox has only a grade of membership to the landmark category. Additionally, this grade of membership to the landmark category depends on the context (Richter & Winter 2014). Winter et al. (2012) strengthened the importance of context-dependent parameters such as mobility, gender, age, education, home town, and other socio-demographic characteristics that influence which object becomes a landmark. This is in line with the early findings of Lynch who already stated in the 1960's that attributes such as age, gender, culture, occupation, temperament, or familiarity of an observer influence the production of the environmental image and the definition of landmarks.

Richter & Winter (2014) summarise the prior results of other researchers and propose the definition of landmarks that is used within this thesis:

> *'Landmarks are geographic objects that structure human mental representations of space' (Richter & Winter 2014, p. 7) and that 'may grab our attention' (Richter & Winter 2014, p. 206).*

A landmark is something that is dependent on people's 'embodied experience and cognitive processing of their living environment' (Richter & Winter 2014, p. 7). A landmark is outstanding because of some attributes or because it generates an experience for an individual structuring the environmental knowledge of a person (Couclelis et al. 1987). Additionally, a landmark contributes to the mental representation of the environment (Richter & Winter 2014).

Characteristics Influencing Salience of Landmarks

The property that turns a conventional geographic object into a landmark is called *landmark salience* (Raubal & Winter 2002, Elias 2003*b*). A landmark should have at least one salient aspect. According to Lynch (1960) the key physical characteristic of a landmark is its singularity that makes this object unique and memorable. Furthermore, Lynch (1960) identifies a clear form, figure-background contrast, and prominence of spatial location as important aspects of an object's salience. A location at decision points or a certain activity attached to an object (e.g. a theatre in a building) may strengthen its importance as a landmark (Lynch 1960).

Inspired by Lynch (1960), further studies regarding the characteristics of landmarks were carried out. Appleyard (1969) determines why people divide urban objects into Lynch's five elements. He discovers the attributes of buildings that capture the attention of people and, therefore, hold a place in their mental representation of a city. He asked a group of inhabitants of the city Ciudad Guayana (Venezuela) about their perception of the city. The inhabitants mentioned a number of buildings, establishments, and other landmarks. These elements were then rated according to a variety of attributes (physical form, visibility attributes, and attributes of use and significance) from which Appleyard (1969) assumed that they might be important for their identification and recall. These ratings were then correlated with the frequencies of element recall in order to identify the relevant attributes. Attributes of physical form (i.e. movement in front of a building, contour, size, and shape of a building as well as its surface), visibility attributes (viewpoint significance and immediacy of a building to the viewing system), attributes of use and significance (such as use intensity and singularity), and other attributes such as recency showed a high influence over recall.

Appleyard (1969) assumes that the relative salience of a building might be more important than any absolute attribute of an object. Therefore, buildings were rated on the basis of their absolute intensity and singularity, both in a local neighbourhood as well as in the whole city. The subsequent regression analysis confirmed that the recall of a building does depend as much on its relation to the context as on any absolute attributes.

A further milestone in landmark research is the characterisation of landmarks proposed by Sorrows & Hirtle (1999). Inspired by Lynch (1960) and Appleyard (1969) their framework defines three key characteristics of an object that influence its salience (Sorrows & Hirtle 1999):

1. **Visual Salience** A geographic object can have salience because of outstanding visual attributes. Visual salience gives information about the visual characteristics of an object in contrast with surrounding objects (e.g. salient shape, colour, or façade area).

2. **Cognitive Salience** An object with an outstanding meaning can have cognitive salience. It may be a landmark because of its typical, but also because of its atypical meaning in the surroundings. The object might have cultural or historical importance or a contrasting content to the surrounding objects.

3. **Structural Salience** A structural salient object is outstanding because of its location in the structure of the environment. Structural landmarks are highly accessible and may have a prominent location (e.g. directly at a decision point).

Burnett et al. (2001) propose alternative characteristics influencing the salience of an object. They suggest *permanence*, *visibility*, *usefulness of location*, *uniqueness*, and *brevity* of a landmark description as the main characteristics of landmarks. Their study focuses on landmarks in terms of usability for car navigation and revealed that salience is dependent on the mentioned characteristics. The characteristics of Burnett et al. (2001) do largely correlate with those of Sorrows & Hirtle (1999). That means, visual salience is equivalent to visibility as structural salience is to usefulness of location (which deals with the location of a landmark in relation to a decision point).

These categories are not mutually exclusive (Duckham et al. 2010). Normally, an object shows more than one characteristic that determines its overall salience as a landmark. Whether an object becomes a landmark is not only affected by exogenous factors but also endogenous factors. Caduff & Timpf (2008) model these factors as a three-valued vector. The components of the vector include exogenous/passive and endogenous/active modes within this model. *Perceptual salience* is the passive mode and defines the potential of a geographic object for acquisition of visual salience. *Cognitive salience*, as the active mode, is triggered by informative cues and provides advance information about a target location. It subsumes endogenous factors that influence the overall salience, which are dependent on the observer's experience and knowledge (Silva et al. 2006). Finally, Caduff & Timpf (2008) introduce *contextual salience* as the third value of the vector. Contextual salience is tightly coupled with modality describing the mode of transportation and task to be performed in the assessment of potential landmarks.

Categorisation of Landmarks

There are a number of ways to categorise landmarks. Possible categorisations are for example according to their location, with regard to a specific route, or according to their spatial extent.

One possible categorisation is in *distant/global* and *local* landmarks (Lynch 1960, Steck & Mallot 2000). Global landmarks are visible from many angles and distances (Lynch 1960) and define a 'global reference frame that does not change when the observer moves a small distance' (Steck & Mallot 2000, p. 69). Global landmarks have some sort of compass function, such as towers, mountain peaks, or skyscrapers. Mobile points (such as the sun) whose motion is slow and regular might be used as a global landmark. In contrast, local landmarks are visible only in restricted areas (Lynch 1960, Steck & Mallot 2000). These are stores, restaurants, metro stations, or signs in an urban environment. Navigating with local landmarks includes a sequence of intermediate goals with local landmarks at theses goals (Steck & Mallot 2000). They are increasingly used for navigation as an observer becomes more familiar with an environment (Lynch 1960). Steck & Mallot (2000) conduct an experiment showing that both, global and local landmarks, are used by travellers. However, some of the participants used only global landmarks while other only used local ones. There were participants who used both types of landmarks. Even though some participants showed a preference for one landmark type, Steck & Mallot (2000) show that the other type was nevertheless present in their memory and available for navigation.

Global and local landmarks are also known as *on-route* (not located at a decision point) and *off-route* (not in the vicinity of the route) landmarks (Lovelace et al. 1999). This categorisation is supplemented by landmarks located at decision points and landmarks at potential decision points (Lovelace et al. 1999). Lovelace et al. (1999) carry out an experiment investigating the use of landmarks in directions for familiar and unfamiliar routes. They show that for familiar routes landmarks at potential decision points are important for the quality of the directions. In addition, they state that for unfamiliar routes landmarks at decision points are most important. Lovelace et al. (1999) explain this difference in landmark type used may stem from experience. Familiar people may remember more landmarks, and, thus, also landmarks independent of decision points, because they had likely used them in the past. For unfamiliar routes the decision points and which way to turn at these points is maybe all that participants can remember after just one exposure (Lovelace et al. 1999). However, Michon & Denis (2001) confirm the clear tendency for landmarks located at decision points. They prove that landmarks are more likely to be mentioned when they are close to a decision point. Further, they find out that a large number of landmarks are mentioned around the starting point of a route and in the vicinity of the destination.

Another possible characterisation is dependent on the spatial extent of a landmark. There are *two-dimensional* landmarks, 'public thoroughfares' (Michon & Denis 2001, p. 295), such as places, streets, and channels and *three-dimensional* geographic objects, such as monuments, buildings, or fountains (Michon & Denis 2001). Michon & Denis (2001) show that in directions of different routes the average number of mentioned landmarks from each category is constant. Overall more three-dimensional landmarks are included in directions (Michon & Denis 2001). Further, they report a difference between women and men. In their experiment women tended to mention more two-dimensional landmarks. Even the route itself or intersections of roads can be a landmark (Klippel & Winter 2005). In contrast, there is no difference between women and men in mentioning three-dimensional landmarks (Michon & Denis 2001).

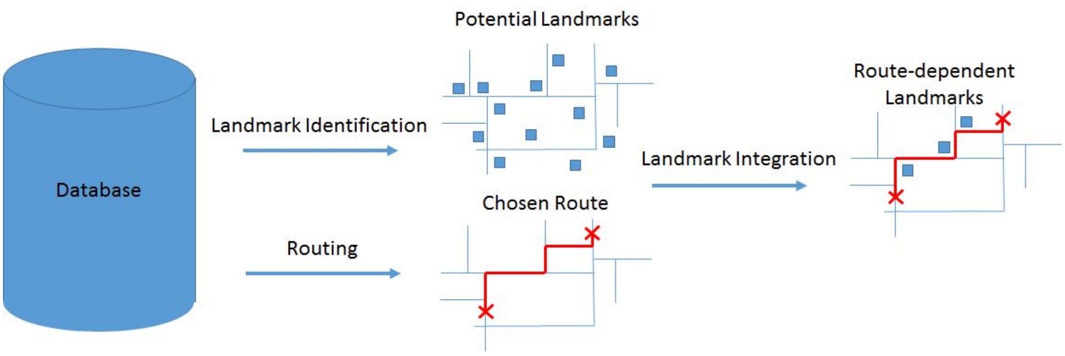


Figure 2.1: Landmark identification and integration (modified from Elias (2003*a*)).

2.2 Modelling of Landmarks for Route Directions

Over the past decades a lot of research was carried out in the field of landmark modelling. There is a distinction between approaches to *landmark identification* and to *landmark integration* (Richter & Winter 2014):

1. Landmark identification concerns the assessment of object salience for navigation and results in a pool of potential landmarks and
2. Landmark integration determines landmarks from the potential landmarks that can be used for a specific route or calculates routes based on landmark information.

Both approaches are important steps in modelling landmarks, but in existing approaches they are performed by different algorithms and research addresses either one or the other (Figure 2.1). The following sections give an overview of the most important approaches.

2.2.1 Landmark Identification

In the research seen here landmark identification considers similar steps (Sadeghian & Kantardzic 2008):

1. specifying a neighbourhood around a potential decision point,
2. identifying objects with outlier characteristics with the help of different methods in the specified neighbourhood, and
3. establishing these salient objects as landmarks.

The following paragraphs elaborate on possible statistical, data mining, hierarchical, and crowdsourcing methods for landmark identification.

Statistical Analysis Methods

Raubal & Winter (2002) propose the first approach towards a formal measure of object salience. They use a weighted sum model based on the characteristics of Sorrows & Hirtle (1999) (Section 2.1.2). Instead of *cognitive salience* they use the term *semantic salience*, which focuses on the meaning of an object.

There are several studies which extend the basic model of Raubal & Winter (2002). Nothegger et al. (2004) extend and test the approach on built-up features, namely façades. They evaluate the concept with human judgement and with real world data. They show that the model from Raubal & Winter (2002) allows for the automatic identification of features which are highly correlated with human choices of landmarks. Further, Winter (2003) includes *advance visibility* of an object in the basic model. He presumes that an object is more suitable as a landmark if it is visible early along a route in contrast to an object that can only be seen at the very last moment. Winter (2003) takes the direction of travel into account to calculate advance visibility of an object. Klippel & Winter (2005) consider positions of point-like objects along a route dependent on the direction of travel. The position of a landmark along the route influences the ease of conceptualising turning actions in route directions and determines the ease of understanding such directions (Richter & Klippel 2007). This inclusion of structural salience in route directions is an approach which may be attached to either landmark identification or integration (Section 2.2.2).

Data Mining and Hierarchical Methods

Raubal & Winter (2002) and other researchers who build on their work, need many different data sources to collect the information for all the attributes (visual, semantic, structural). They use data sources such as digital city maps, rectified geo-referenced images, and navigation graphs for the actual means of travel. Elias (2003*b*, 2006) explicitly identifies this time-consuming and expensive data collection process as the weak point of landmark identification approaches. Therefore, she proposes to use existing spatial databases instead of manual collection methods. She focuses on point-like buildings as landmark candidates and uses spatial attributes from topographic and cadastral datasets to automatically extract landmarks using data mining methods. She specifies a neighbourhood to investigate dependent on the density of the buildings around a decision point. In case of a low density of buildings (i.e. areas with open spaces and parks) she selects a larger neighbourhood for her analysis (Elias 2003*b*). She uses a small neighbourhood if there are a lot of buildings available. The buildings within the specified neighbourhood have attributes that either refer to land use attributes or to geometrical ones (size of building, orientation to road, number of corners, ...). Elias (2003*b*, 2006) aims to identify objects with outlier attributes to determine buildings which are unique in the specified neighbourhood. She uses an adaptation of a classical decision tree machine learning approach based on the entropy principle, namely ID3 (Quinlan 1986) to identify such buildings.

Winter et al. (2008) propose a computational model for the generation of a hierarchy of landmarks, combining the approaches of Raubal & Winter (2002) and Elias (2006). The hierarchy presents a ranking order for landmarks based on their individual saliences. The landmark is seen as an anchor point of the region in which the landmark is the most prominent object. Neighbouring landmarks are compared by prominence and only the most salient ones are taken into the next level of hierarchy. This results in a classified hierarchy (Winter et al. 2008) which is usable for various tasks (e.g. for destination descriptions (Tomko & Winter 2009)).

Volunteered Geographic Information (VGI) and Crowdsourcing Methods

The idea of citizen involvement in carrying out various activities relating to geo information systems (See et al. 2016) emerged from different disciplines in past years: e.g. wikification of geospatial information for the wide masses (Boulos 2005), crowdsourcing (Howe 2006), user-generated content (Krumm et al. 2008), and VGI (Goodchild 2007), to name just a few.

Tezuka & Tanaka already recognised in 2005 that the internet provides a rich source of spatial information. They investigate how geographic objects are expressed by humans and extend existing methods of text mining in such a way that spatial context is considered. They show that using these methods improves the precision of extracting landmarks from web documents.

Quesnot & Roche (2014) argue that social location sharing datasets are a reliable data source to retrieve the semantic salience of landmarks. Richter (2017) highlights the increasing availability of user-generated content with geographic components which could be exploited for identifying landmarks. A number of VGI and crowdsourcing methods deal with the use of OpenStreetMap (OSM) data which are made available via the web. Richter & Winter (2011) report on integrating landmarks in OSM and demonstrate the advantages of user-generated content for extracting semantic information. Another approach assesses the suitability of an object as a landmark using a landmark index based on attribute values of buildings extracted from OSM (Nuhn et al. 2012). The data are complemented by 3D city models (e.g. for height information). Nuhn et al. (2012) investigate if an attribute value of a building is salient or differs from the attribute values of the surrounding buildings. They obtain a landmark index by adding the salience of the individual attributes and dividing the sum by the number of attributes. A building is classified as a landmark if the landmark index exceeds a predefined threshold.

Wolfensberger & Richter (2015) propose another crowdsourced approach, which uses OSM data. They introduce a mobile application, which enables a user-generated collection of landmarks. Based on a photo taken by a smartphone the application calculates and ranks potential landmark candidates taking the current visibility area into account. The most probable landmark candidate is calculated by allocating measurable attributes to characteristics of Sorrows & Hirtle (1999) (Section 2.1.2). The landmarks are presented to the user who then may choose the intended one.

Kattenbeck (2015, 2016) addresses the lack of available data sources within his work and identifies the use of crowdsourced data acquisition approaches to overcome this problem. He proposes an empirically validated model and an approach to survey-based assessment of the salience of an object. In his work, he uses a structural equation model and incorporates the results of prior studies and features which are important salience indicators. The model was empirically tested in the framework of a large scale in-situ experiment. Kattenbeck (2015, 2016) reveals a high impact of visual salience on visibility in advance, which, in turn, had an influence on structural salience. Another interesting finding is that he identifies emotional salience and familiarity as two possibly missing subdimensions of salience of a geographic object. Kattenbeck's (2016) model was transferred to another city to assess the invariance with respect to the environment, its objects, and the observers (Kattenbeck et al. 2018). The results showed that the relationships between the subdimensions of salience does not differ significantly in another environment. Hence, authors state that the model can be used to calculate salience across different environments.

2.2.2 Landmark Integration

While there has been a lot of research on landmark identification, there is only little research on landmark integration. There are two main directions: firstly, determining landmarks for a specific route and secondly, calculating optimal routes based on identified landmarks.

Determination of Landmarks for a Specific Route

Tomko (2004) proposed to assess the suitability of data from the world wide web to provide information for pedestrian navigation even before all the definitions of VGI and crowdsourcing became popular (Section 2.2.1). His experiment provides landmark based directions along a path generated by a web service, which is tested and evaluated with human subjects. The results show that, already at this early stage of internet development, the web is capable to provide elements that can complement and enhance route directions.

Klippel & Winter (2005) take into account the location of a landmark relative to a turn at a decision point and the kind of wayfinding action that needs to be performed (Section 2.2.1). Apart from advance visibility (Winter 2003) their model considers the configuration of the street network as well as the route along the network. The result is a mathematical measure that describes the ideal position of a landmark at an intersection. Röser et al. (2012) examine the different landmark positions in two experiments. First, from a bird's eye perspective and second, from an egocentric perspective. Their results provide evidence to support the assumption of Klippel & Winter (2005) that visibility and structural salience are interdependent empirically.

Richter (Richter 2007, Richter & Klippel 2007, Richter 2008) integrates landmarks that allow for the easy conceptualisation of spatial situations. He reflects on how landmarks are referred to in human route directions and employed concepts such as 'before' or 'after a turn' (Klippel & Winter 2005). Richter exploits ordering information to determine a landmark's relative location to a turn using point-like as well as linear and areal landmarks in different spatial situations. The concept is implemented and tested in a system called GUARD (generation of unambiguous adapted route directions).

Winter et al. (2009) and Duckham et al. (2010) propose a completely different approach to integrate landmarks in route directions. While other methods are based on visual or geometric characteristics of individual objects, their approach relies solely on information about the types of landmarks. Sorrows & Hirtle (1999) introduce the concept of prototypicality which goes back to the work of Rosch et al. (1976) and describes how typically a landmark represents a category. The aspect of prototypicality plays an important role in the model of Duckham et al. (2010). They develop a weighting system that assigns weights to Point of Interest (POI). The weights are dependent on the suitability of a typical POI category as a landmark and the likeliness that a POI category is typical. Properties of categories of POIs (e.g. ubiquity, length of description, permanence, ...), such as might be found in a directory service such as the Yellow Pages (e.g. Hotels, Restaurants, Parks, Museums, etc.) are considered. Duckham et al. (2010) assign the weights to the POIs and use a standard algorithm (e.g. Dijkstra (Dijkstra 1959) or A* (Hart et al. 1968)) to generate a route. They select the POI which coincides with a decision point on the route and that shows the highest weight and include it in the route directions. This approach is implemented in the *WhereIS* route service (Sensis 2017) using categories from the Yellow Pages.

Currently, there is a focus on determining global landmarks for a specific route (Wenig et al. 2017, Credé et al. 2017). Wenig et al. (2017) present Pharos, a new system to include global landmarks in route directions. They show that the visibility of global landmarks can be derived from existing and publicly available geotagged images which is an advantage over hard to select local landmarks. They demonstrate that participants navigate more confidently and build a more accurate cognitive map by including global landmarks.

Other approaches focus on the modelling of landmark-based navigation directions from open source data. Dräger & Koller (2012) present an approach for car navigation that relies exclusively on OSM data. Their system chooses appropriate landmarks at decision points and includes them in route directions. Rousell et al. (2015), Rousell & Zipf (2017) propose an approach to integrate landmarks in route directions for pedestrians based on identifying the contextual type of an object from OSM data. Additionally, they consider geometric calculations in relation to a decision point. The implementation shows that suitable landmarks can be successfully extracted and integrated into route directions for a specific route from the OSM dataset.

Calculation of Optimal Routes Based on Landmark Information

A second way of landmark integration, besides determining landmarks for a specific route, is calculating optimal routes based on landmark information. Caduff & Timpf (2005*a*,*b*) propose the Landmark-Spider-Algorithm to calculate the clearest route in terms of landmarks. It navigates a traveller along a route with selected landmarks used to give route directions at every decision point. Authors select landmarks based on the salience of spatial objects and on distance and direction of the traveller with respect to the objects. They present the results of this algorithm in a spatio-analogical way which supports wayfinding decisions. Rüetschi et al. (2006) propose another approach to incorporate landmarks in route generation algorithms. They use landmarks as parts of route directions and map them to sets of edges in a street network. They build auxiliary graphs in such a way that a standard shortest path algorithm can be used to find an optimal route.

Other approaches use a modified Dijkstra algorithm (Dijkstra 1959) to determine optimal routes regarding landmark information (Elias & Sester 2006, Chandrasekara et al. 2016). Elias & Sester (2006) use point-like buildings as landmarks which are identified using the approach of Elias (2006). They apply the Dijkstra algorithm to identify an optimal route based on landmark quality. Their idea is to describe the cognitive complexity of a route in relation to the quality of landmarks and the corresponding route directions. Elias & Sester (2006) adapt weights according to the visibility, usefulness of location, uniqueness, permanence, and brevity of a landmark description. This leads to the identification of an optimal route in terms of cognitive load to remember and follow the route direction. A recent approach makes also use of the Dijkstra algorithm. Chandrasekara et al. (2016) consider besides distance information the strength of landmarks along a route. They derive the strength of landmarks based on landmark density along an edge and their significance for navigation. To determine landmark salience horizontal spread, height, and the visibility of landmarks at different times of the day as well as the social/cultural salience is considered. The approach is implemented using OSM data and verified and tested in Sri Lanka.

2.3 Towards Personalised Landmarks

Landmark salience is not the same for every person and dependent on parameters such as age, gender, education, or familiarity of the traveller with an environment (Lynch 1960, Winter et al. 2012). A large body of research deals with the adaptation of the content and appearance of maps based on user preferences (e.g. Sarjakoski et al. (2007), Sarjakoski & Sarjakoski (2008), Reichenbacher (2007), Wiebrock (2011)). For example, the knowledge-based system by Sarjakoski et al. (2007) considers aspects such as the time (e.g. seasons or time of the day), the use case for which a map is needed (e.g. outdoor, cycling, or emergency), or the user's age group. However, they discuss no other parameters about the traveller's knowledge or experience and especially no landmarks are considered. Burnett et al. (2001) are one of the first who show that travellers being familiar with an environment choose other landmarks for route directions than people unfamiliar with an environment. In an experiment, two conditions were adopted, whereby participants provided route directions based on either long-term experience or single experience. Participants with single experience had no prior experience of the route whereas long-time experienced participants had lived and/or worked in the area for at least five years. The study shows that participants with single experience of an environment refer more to general salient objects related to the street (e.g. pedestrian lights, churches, and petrol stations) whereas people with long-term experience refer to more specific things (e.g. specific restaurants, bingo halls, or toy shops).

Winter et al. (2005) propose an approach to adapt Raubal & Winter's (2002) model to different user-contexts. They include context information by modelling weights for the salience measures (Section 2.2.1). In addition, they investigate the proposed method in a thorough human subject test. They find evidence that the variation of the context changes the selection of the landmark. However, their work focuses on weights based on different contexts (here, the time of the day). Apart from gender differences in weighting landmarks by day and by night, no other personal attributes are treated. Although the familiarity with the environment was collected from test persons on a simple binary scale this attribute is not further evaluated.

The crowdsourced data acquisition approach of Kattenbeck (2016) (Section 2.2.1) includes, amongst others, questions on demographic data. This includes e.g. the background of a traveller and the knowledge about a place. Kattenbeck (2016) assumes that 'knowledge about a local neighbourhood may have an effect on several dimensions of salience' [p.91]. However, this information was captured to minimise the bias in salience estimations of objects but was not further evaluated.

More recent studies show that famous buildings are more easily recognised than unfamiliar ones (Hamburger & Röser 2014). These differences cannot be explained by visual characteristics, because authors choose comparable visual salient buildings for the experiment. These results provide empirical evidence for the assumption that familiarity or cognitive salience (Caduff & Timpf 2008) is relevant for overall salience of a landmark. Based on these findings Quesnot & Roche (2015) assume that travellers unfamiliar with an environment prefer different landmarks than travellers who know the area well. They confirm this assumption and show that persons familiar with an environment prefer landmarks with cognitive or semantic salience respectively. In contrast to that, for unfamiliar people visual salience is more important than semantic salience (Quesnot & Roche 2015). Recently, Sameer & Bhushan (2017) investigate the effect of familiarity and degree of recognition as important components of cognitive salience. They draw the same conclusion as the other researchers and indicate that familiar buildings are better landmarks than unfamiliar ones.

Current work investigates differences between classical route directions and modified route directions: firstly, non-personalised modified directions including irrelevant information about landmarks, and secondly modified directions including information of *personal interests* associated with landmarks (Gramann et al. 2017, Wunderlich & Gramann 2018). Participants of a study provided individual preferences such as taste of food, music, or favourite animals. Gramann et al. (2017), Wunderlich & Gramann (2018) used this information to modify the directions and include this *personal interests*. The authors confirmed enhanced spatial memory performance and landmark recognition for the modified route directions without further differentiating between personalised and non-personalised directions. This means, the modified directions with *personal interests* did not perform better than the modified non-personalised directions.

Meng (2005) shows that the usability of egocentric mobile maps is dependent on subjective parameters, such as e.g. the users emotion (e.g. joyfulness or irritation) during map interaction. Schroder et al. (2011) highlight the importance of emotions towards features, although they state that emotion is an aspect of landmark salience which is difficult to model. Balaban et al. (2014) focus also on emotion and especially on affect (Balaban et al. 2017). They introduce a new landmark salience category: *emotional landmark salience*. In their studies they consider the mood condition (positive, negative, neutral). An experiment revealed that participants show higher wayfinding performance for negatively laden landmarks than for positively laden landmarks and a higher performance for positive landmarks than for neutral landmarks. In addition, negative landmarks are better remembered than positive and neutral landmarks because recognition performance hardly decreased over time for these landmarks. Furthermore, Palmiero & Piccardi (2017) show that both, positively and negatively laden landmarks, equally support path learning and, therefore, influence the acquisition of spatial knowledge. They show that positive emotional landmarks improved the reproduction of a path on the map compared to negatively or neutrally laden landmarks. Ruotolo et al. (2018) support the finding that emotional factors influence perception and memorisation of spatial dimensions. They show that positions of landmarks along a route with neutral or negative values are remembered less accurately than the positions of positive landmarks.

Götze & Boye (2013, 2016) propose to learn individual salience models for landmarks that are referred to in route directions. They model every landmark a person refers to as a feature vector including several attributes (e.g. distance and angle to a landmark as well as name and type extracted from OSM data). Then, they calculate the salience of a landmark as a weighted sum of the elements of the feature vector. They derive a person's salience model that calculates which object is most suitable to be used in route directions. The evaluation of their models show promising results, since their model was often able to predict the landmark chosen by a person.

There is a new direction or research aiming at exploring interesting landmark recommendations based on geo-tagged photos (Shi et al. 2011, Chen et al. 2013, Han & Lee 2015). Shi et al. (2011) base their recommendation on the assumption that a traveller in a new city may like landmarks that are already favoured by other users with similar landmark visiting experiences in other cities in the past. A similar approach is followed by Han & Lee (2015) who compute the significance of an object for a traveller based on their trips' spatial and temporal properties. Based on travel trajectory history they generate clusters of landmarks with similar or related themes for recommendations.

Personalisation is 'the process of making something suitable for the needs of a particular person' (Cambridge Dictionary 2019). Nuhn & Timpf (2016) present the first ideas of identifying suitable landmarks for the needs of specific persons with the help of a multidimensional model for personalised landmarks. In Nuhn & Timpf (2017*b*) they identify personal dimensions of landmarks as a basis for such a multidimensional model and their attributes. Nuhn & Timpf (2017*a*,*c*) propose a conceptual framework for a multidimensional model for personalised landmarks that integrates three dimensions: a dimension describing the landmark, an environmental dimension, and a personal dimension. They identify and discuss attributes as well as attribute values for each of the dimensions and develop salience measures for them (Nuhn & Timpf 2017*a*). Nuhn & Timpf (2018) include the personal dimensions *prior spatial knowledge*, *personal interests*, and *personal background* in their multidimensional model. They present a conceptual model without the empirical evidence that the addition of personal dimensions to a landmark salience model may result in more identified landmarks than a conventional model without personal dimensions.

.

2.4 Implications for Modelling Personalised Landmarks

In this chapter we investigated cognitive aspects of human wayfinding and identified piloting between landmarks as an efficient means of travelling to familiar and novel destinations. Piloting between landmarks to accomplish a wayfinding task is the means dealt with in this thesis. We identified spatial knowledge as an important information to enable people to find routes in geographic space and we investigated the three stages: landmark, route, and survey knowledge (Siegel & White 1975). We will build on these three stages for the modelling of the personal dimension *prior spatial knowledge*. Furthermore, we showed that even the shortest route directions that a human communicates include landmarks (Lovelace et al. 1999) and we evaluated definitions and characteristics influencing the salience of these landmarks. Richter & Winter (2014) summarised the prior results of other researchers and proposed a definition of landmarks. Sorrows & Hirtle (1999) defined three key characteristics of an object that influence its salience, which are the basis for our landmark dimensions: the visual, the semantic, and the structural dimension.

We showed that over the past decades a lot of research was carried out in the field of landmark modelling. There is a distinction between the approaches to landmark identification and those to landmark integration (Richter & Winter 2014). We gave an overview of the most important approaches and found out that landmark identification and integration are performed by different algorithms and that research addresses either the one or the other respectively. We base this work on landmark identification models and investigate amongst other models the existing weighted sum model for landmark identification proposed by Raubal & Winter (2002). Furthermore, we intend to use a decision tree model which is a machine learning approach already used for landmark identification in the past (Elias 2006).

We discussed prior work concerning personalised landmarks and it actually reveals that landmark salience is not the same for every person but dependent on several parameters (Lynch 1960, Winter et al. 2012). Based on the landmark definition of Richter & Winter (2014) and the definition of personalisation (Cambridge Dictionary 2019) we define a *personalised landmark* as follows:

Personalised 'landmarks are geographic objects that structure human mental representations of space' (Richter & Winter 2014, p. 7), that 'may grab our attention' (Richter & Winter 2014, p. 206), and that are suitable for our needs.

We revealed familiarity as one important parameter resulting in different landmark preferences (Hamburger & Röser 2014, Quesnot & Roche 2015, Sameer & Bhushan 2017). Therefore, we assume that the suitability of a geographic object as a personalised landmark is dependent on *prior spatial knowledge*. We found first studies investigating *personal interests* and personalised landmarks (Gramann et al. 2017, Wunderlich & Gramann 2018). These studies do not suggest that there might be benefits of considering *personal interests* compared to other information about landmarks. However, we consider this dimension in our personalised landmark identification models to confirm or reject these findings.

We investigated in this chapter first efforts to identify personalised landmarks. Winter et al. (2005) adapted the model of Raubal & Winter (2002) focusing on weights based on different contexts. Apart from gender differences in weighting landmarks by day and by night, no other personal attributes are treated. We will build on the approach of Raubal & Winter (2002) focusing on weights based on personal dimensions. Nuhn & Timpf (2018) include the personal dimensions *prior spatial knowledge*, *personal interests*, and *personal background* in a multidimensional model to identify personalised landmarks. They present their conceptual model without the empirical evidence that the addition of personal dimensions to a landmark salience model may result in more identified landmarks than the conventional model without personal dimensions. This empirical evidence is still missing.

We assume that the collection of personal data is the highest effort for the identification of personalised landmarks. Therefore, we need to make sure that the data collection effort is justified relative to the benefits that can be achieved through the provision of personalised landmarks. However, so far, there is no computational landmark identification model available that includes personal dimensions. Thus, there has been no comparison possible between a *conventional* and a *personalised landmark identification model*. This means it is an open question whether a personalised landmark identification model incorporating *prior spatial knowledge* and *personal interests* identifies more landmarks selected by humans than a conventional, non-personalised model. We intend to develop models for personalised landmark identification and compare them with conventional, non-personalised models. For this calculation several mathematical models and analysis methods are possible and we will investigate them in the following chapter.

Chapter 3

Mathematical Models and Analysis Methods

The aim of our landmark identification models is to find all objects at a decision point that are able to be a (personalised) landmark. Therefore, we need to use models that are able to do so. A landmark might be identified by either calculating an overall salience measure or by classifying objects as *landmark* (LM) and *not a landmark* (NAL). In this chapter (Section 3.1) we investigate three models based on theory: a weighted sum model (wSm), a weighted product model (wPm), and a decision flow chart (dFc). In addition, we investigate a decision tree model (dTm) which is an approach in the field of machine learning. The wSm and the wPm calculate an overall measure of landmark salience for an object, whereas the dTm and the dFc classify objects as LMs and, in the case of the dTm, NALs.

We intend to build conventional and personalised landmark identification models (Section 4.3). The machine learning models both, conventional and personalised, learn their behaviour from examples and are able to generalise after learning. For this to happen, the model needs to learn its model parameters from data via a process called *training*. The resulting models are able to identify whether an object of a new unseen dataset is a LM or a NAL (*testing*). In Section 3.2.1 we discuss the traditional machine learning approach for training and testing. The conventional models based on theory have no unknown model parameters, whereas the model parameters of the personalised models that are also based on theory need to be identified. Inspired by the traditional machine learning approach, we investigate in Section 3.2.2 methods to 'train' and test these models based on theory.

The training results in conventional and personalised landmark identification models ready to identify landmarks of a new unseen dataset. We investigate methods to analyse the trained models and their results. This includes methods for sensitivity analysis (Section 3.3) and for the comparison of the model results (Section 3.4). We close this chapter with an outlook on the study setup of this thesis (Section 3.5).

3.1 Mathematical models

We investigate a weighted sum model (wSm), a weighted product model (wPm), and a decision flow chart (dFc) inspired by theoretical considerations. In addition, we investigate a decision tree model (dTm), which is an approach in the field of machine learning.

3.1.1 Models based on Theory

In this section we investigate models inspired by theoretical considerations. In contrast to machine learning models, these models do not learn from data but are based on predefined established models and algorithms (Srinivasan 2016).

Weighted Sum Model (wSm)

A widely used model is the wSm (Triantaphyllou 2000). It applies the additive utility hypothesis, which 'implies that the overall value of every alternative is equivalent to the products' total sum' (Kolios et al. 2016, p. 5). The wSm is best suited for problems with attributes of the same units. In case of varying units (e.g. quantitative and qualitative attribute values) normalisation schemes should be employed (Kolios et al. 2016). If there are m alternatives and n attributes, then the best alternative is obtained with the following formula (Fishburn 1967):

$$A_{wSm} = max \sum_{j=1}^{n} a_{ij} * wj, \qquad for\ i = 1, 2, 3, ..., m. \tag{3.1}$$

A_{wSm}: wSm score of the best alternative
aij: score of the i-th alternative with respect to the j-th attribute
wj: weight for the j-th attribute
n: number of attributes
m: number of alternatives

Assume that you want to choose the best alternative among A_1, A_2, and A_3. The attributes are a_1, a_2, and a_3. Table 3.1 shows example aij values and weights wj .

Table 3.1: Example of alternatives and attributes for Weighted Sum Model (wSm) and Weighted Product Model (wPm).

	Attributes		
Alternatives	a_1	a_2	a_3
Wj	0.3	0.3	0.4
A_1	25	50	100
A_2	50	75	100
A_3	100	50	50

When the Formula 3.1 with the data delivers: $A_1 = 62.5$, $A_2 = 77.5$, and $A_3 = 65$. Based on these results A_2 is the best choice, because the value of A_2 is the highest of the values of the alternatives.

Weighted Product Model (wPm)

Bridgman (1922) introduces the wPm. It is an alternative to the wSm but is not widely utilised (Yoon & Hwang 1995). The main difference to the wSm is that a product is applied in the model instead of a sum. Because the attributes are connected by multiplication normalisation schemes are not needed (Azar 2000) in case of varying units. The wPm sets the weights as exponents of each attribute value. The formula for the best alternative is as follows (Budiharjo & Abulwafa 2017):

$$A_{wPm} = max \prod_{j=1}^{n} a_{ij}^{wj}, \qquad for\ i = 1, 2, 3, ..., m. \tag{3.2}$$

A_{wPm}: wPm score of the best alternative
aij: score of the i-th alternative with respect to the j-th attribute
wj: weight for the j-th attribute
n: number of attributes
m: number of alternatives

The Formula 3.2 with the numbers in Table 3.1 delivers $A_1 = 53.59$, $A_2 = 74.51$, and $A_3 = 61.56$. Thus, the wPm produces the same result as the wSm. A_2 remains the best choice, because the value is the highest one of the alternatives.

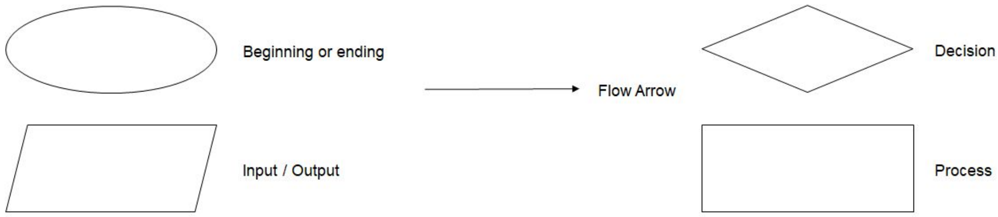


Figure 3.1: Flowchart symbols (modified from Myler (1998)).

Decision Flow Chart (dFc)

There is a large body of knowledge available on landmarks, landmark salience, and dimensions influencing the landmarkness of an object (Section 2). In order to take this knowledge into account we intend to build a decision flow chart to depict assumed interdependencies between landmark as well as personal dimensions and the landmark salience of an object. There is a long tradition of using diagrams to represent decision problems. Gilbreth & Gilbreth (1921) introduce the first method for documenting processes. This fundamental work serves as a basis for a standard for *flow process charts* (ASME 1947). A flowchart is a graphical representation. There are flowchart symbols provided by the ISO in 1970 and revised in 1985 (ISO 1985). Figure 3.1 shows some of the common flowchart symbols.

Fryman (2002) differentiates types of flowcharts including decision flowcharts. Building a decision flowchart consists of several steps (Fryman 2002, LucidChart 2018, Graham 2004):

1. Defining the area of focus.
2. Conducting a thorough literature research.
3. Identifying the steps in chronological order.
4. Generating hypotheses in order to identify decisions, processes, inputs, and outputs.
5. Establishing decision rules for accepting or rejecting hypotheses.
6. Drawing the flowchart.
7. Confirming the flowchart with validation data.

Flowcharts flow from left to right and top to bottom (Myler 1998). Decisions may be multiple choice or two-way decisions (Fryman 2002). It depends on the application, which symbols and decision types are included in the flowchart.

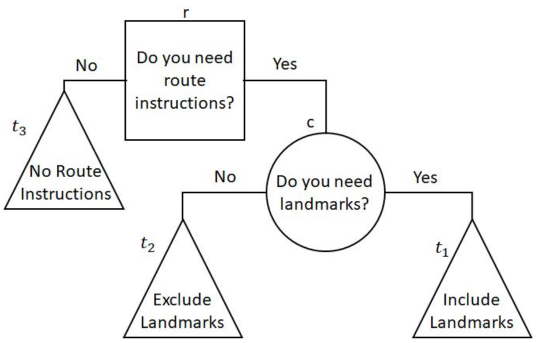


Figure 3.2: A sample decision tree (modified from Kamiński et al. (2018)).

3.1.2 Machine Learning Model

Machine learning is a research field that gives a model the ability to learn its behaviour from data (Samuel 1959). There are numerous machine learning methods. We intend to apply a model similar to the dFc. One popular approach of machine learning for classification is a decision tree model (dTm) (Rokach & Maimon 2005). These are flow chart-like structures (Gupta et al. 2017), whose main difference to a dFc is that they do not consider processes and concentrate only on decisions and their results. The idea is to break up a complex decision problem into a number of simpler decisions (Safavian & Landgrebe 1991). After every decision another decision follows until a conclusion about the class of the object is reached (Tan et al. 2006). This technique is used in applied fields such as finance, marketing, engineering, and medicine (Rokach & Maimon 2015). Hyafil & Rivest (1976) state that decision trees are np-complete because of the large effort put into finding efficient optimal algorithms for constructing optimal binary trees.

Decision trees are generated from training sets of the form:

$$(x, Y) = (x_1, x_2, x_3, ..., x_n, Y) \tag{3.3}$$

Y is the dependent target variable for the classification. The target variable can take at least two values (e.g. LM and NAL). The vector x has attributes $i = 1....n$ that are used for the classification (Safavian & Landgrebe 1991). The tree is constructed using a directed graph with nodes V and edges (branches) E: $G = (V, E), E \subset V^2$. The set of nodes V consists of three disjoint sets $V = R \cup C \cup T$ (Kamiński et al. 2018).

- the root node (R) is the initial state of the decision tree (Apté & Weiss 1997), it has no incoming edges and zero or more outgoing edges,

- a number of internal nodes (C) with one incoming edge and two or more outgoing edges (Tan et al. 2006), and
- a terminal or leaf node (T), which is the end node with one incoming edge and no outgoing edges (Tan et al. 2006).

In Figure 3.2 there are nodes $R = r$, $C = c$, and $T = t_1, t_2, t_3$ and edges $E = (r, c), (r, t_3), (c, t_1), (c, t_2)$ (Kamiński et al. 2018).

Generally, there are many decision trees that can be built from a given dataset (Tan et al. 2006). Finding an optimal tree is feasible only in small problems (Rokach & Maimon 2005). Efficient algorithms are developed to induce a reasonably accurate decision tree in a reasonable amount of time (Tan et al. 2006). There are four most widely used decision tree models (Lin et al. 2006, Song & Lu 2015): Classification and Regression Trees (CART) (Breiman et al. 1984), C4.5 (Quinlan 2014), CHAID (Chi-Squared Automatic Interaction Detection) (Kass 1980), and QUEST (Quick Unbiased, Efficient, Statistical Tree) (Loh & Shih 1997).

Most algorithms generate the tree in a top-down approach (Apté & Weiss 1997) meaning the number of attributes becomes smaller as the tree is traversed (Tan et al. 2006). The algorithm starts tree growing with the entire dataset in the root node (Ture et al. 2009). Each iteration of the algorithm splits each node into two or more internal nodes according to a certain discrete function (Rokach & Maimon 2005). The goal is to produce data subsets which are as homogeneous as possible with regard to the target variable Y (Breiman et al. 1984). There are a number of different functions available for splitting, such as impurity based criteria. One widely used function is the gini-index (Breiman et al. 1984, Gelfand et al. 1989), which is the probability of obtaining two different outputs and calculates as follows (Breiman et al. 1984, Gelfand et al. 1989):

$$gini = 1 - \sum_{i=1}^{J} p_i 2, \qquad for\, i = 1, 2. \tag{3.4}$$

J: number of classes
pi: the fraction of objects labelled with class i in the dataset

Further possible functions are entropy or information gain to construct a decision tree (Quinlan 1986).

Based on the number of edges at the nodes, decision trees are divided in binary and non-binary trees. Most decision tree induction algorithms apply the splitting function to one attribute at a time (Tan et al. 2006). There are multivariate linear decision trees that use multiple attributes for the splitting conditions in the internal nodes (Brodley & Utgoff 1995, Heath et al. 1993, Breiman et al. 1984). There are solutions dealing with nominal, ordinal, or continuous values (Tan et al. 2006). However, finding an optimal multivariate linear split is more difficult than finding the optimal univariate split and for some feature evaluation rules even intractable (Murthy 1998).

The splitting process results in fully grown trees until a stopping criteria is reached. One problem of fully grown trees are that they are likely to overfit the data (Bramer 2007, Dietterich 1995). A decision tree overfits the training data if the tree depends too much on irrelevant attributes of the training set. The result is that its performance is poor on unseen data (Bramer 2007). *Pruning* is carried out to reduce the size of a decision tree (Tan et al. 2006, Mingers 1989). There are various methods for decision tree pruning. Generally, there is a distinction between pre-pruning and post-pruning (Fürnkranz & Widmer 1994). Post-pruning means that the decision tree is generalised after the growing phase. Popular post-pruning algorithms are reduced error pruning (Brunk & Pazzani 1991) or cost-complexity pruning (Bradford et al. 1998). Pre-pruning is applied during decision tree growing and uses some sort of stopping criteria for the model parameters (e.g. depth of a decision tree, minimum samples in a leaf) or condition related criteria (Quinlan 1990, Fürnkranz 1994*b*). Pre-pruning is very efficient and less computationally expensive as post pruning but sometimes post pruning is more accurate (Fürnkranz 1994*a*).

However, there are a lot of decision tree models available for different applications. It depends on the target variable, the values of the attributes, and the general goal, which decision tree model is the most suitable one. We investigate in Section 4.3.7 which one is the most suitable one to identify (personalised) landmarks.

3.2 Model Training, Validating, and Testing

The machine learning models learn their model parameters from *training sets* of the form as shown in Formula 3.3 via training. A training set includes objects with attributes and a target variable whose value is known, i.e. whether an object is a LM or a NAL. After the model training is complete, the model is used for testing, i.e. to identify landmarks on a *test set*. In section 3.2.1 we describe the traditional machine learning approach for model training and testing. The conventional models based on theory do not have any unknown model parameters. However, we need to identify the model parameters of the personalised models based on theory. These models based on theory differ from machine learning models because they are explicitly predefined models and are not learned from any data. However, we decide to pursue a comparable approach to 'train' the models based on theory and adjust the traditional machine learning approach (Section 3.2.2). In Section 3.2.3 we investigate methods how to split a dataset in training and test set.

3.2.1 Traditional Machine Learning Approach

Figure 3.3 shows the traditional machine learning approach for training and testing. The initial dataset is divided into a training set and a test set (Section 3.2.3). The machine learning models learn their behaviour from the objects of the training set and are able to generalise after learning on new unseen data of the test set. The training set includes LMs and NALs and the test set only includes LMs since we are only interested in the identification of landmarks. During the training the test set is entirely separate, *locked away*, and only employed after all model training is completed (Russell & Norvig 2016).

The first step of the training is to feed the model with data from the training set for whose objects it is known whether an object is a LM or a NAL (Figure 3.3). A useful practice to find the optimal model parameters is a grid-search (Chicco 2017). 'Grid Search is the process of scanning data to configure the optimal parameters for a given model' (Reyhana et al. 2018, p.98). For each combination of model parameters of the grid-search we build a model with the goal of identifying the best one (Cambridge Coding Academy 2019). A complete grid-search might be time-consuming. Therefore, Hsu et al. (2016) recommend a two-step approach: first, a coarse grid-search, and after identifying a good region on this grid, a finer grid-search on that particular region.

However, scanning through all possible model combinations, building models, and evaluating them on the test set will provide the combination of model parameters that performs best, but these parameters might not generalise well on new unseen data (Cambridge Coding Academy 2019). A solution for this problem is k-fold cross-validation (Stone 1974, Geisser 1975). For each combination of model parameters of the grid-search the training set is splitted into k subsets (folds) (Figure 3.3). Since the training set includes LMs and NALs, the k-folds also include both. One of these folds is called *validation fold* and the other k-1 folds are generally called *training folds* (Russell & Norvig 2016). We use the training folds to create the model. Subsequently, we use the created model to identify the LMs and NALs of the validation fold. Thus, the validation fold is used to evaluate the model in order to get an early estimate of the model's performance during training and without using the *locked away* test set.

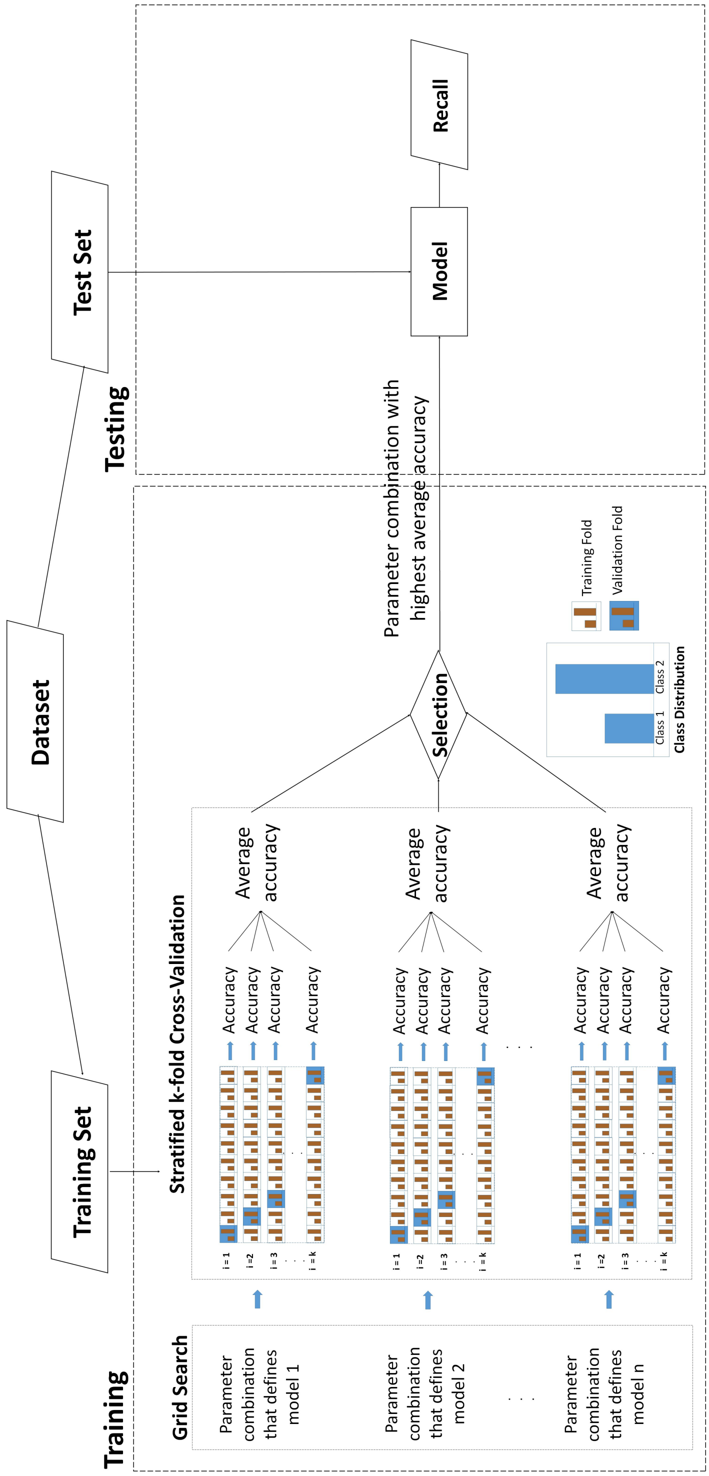


Figure 3.3: Traditional machine learning approach.

We derive the *accuracy* of the model from a *confusion matrix* (Hay 1988) (Table 3.2) that includes information on:

- *Identified LMs*: The object is a LM and is identified as a LM.
- *Identified NALs*: The object is a NAL and is identified as a NAL.
- *Unidentified NALs*: The object is a NAL but is identified as a LM.
- *Unidentified LMs*: The object is a LM but is identified as a NAL.

The formula to calculate the *accuracy* of our machine learning model is as follows:

$$accuracy = \frac{Identified\ LMs + Identified\ NALs}{all\ Objects} \tag{3.5}$$

$$\begin{aligned} all\ Objects = {} & Identified\ LMs + Unidentified\ LMs \\ & + Identified\ NALs + Unidentified\ NALs \end{aligned}$$

Cross-validation uses each fold of the training set only once as a validation fold to calculate the accuracy for a model trained on the other k-1 folds (Kohavi 1995). This process is repeated k times and results in k cross-validation accuracies. These k accuracies are averaged to one k-fold cross-validation accuracy to give an indication on the models performance (Figure 3.3). This can be done for several models with different model parameter combinations, then, the set of model parameters that defines the model achieving the highest average accuracy is selected (Schaffer 1993).

A key issue of k-fold cross-validation is the number of folds. There are many empirical studies verifying that a reliable estimate can be obtained with k = 10 for a dataset with a sample size greater than 100 (Borra & Di Ciaccio 2010). There are a number of ways for 10-fold cross-validation - following Kohavi (1995) we use *stratified* cross-validation. It divides the dataset in disjoint folds with equal class distributions and is preferable both, in terms of bias and variance, compared to regular cross-validation (Kohavi 1995). The folds of stratified cross-validation contain approximately the same ratio of classes as the original dataset (Figure 3.3). Usually, the target values are used as classes for stratified cross-validation in the traditional machine learning approach (LMs and NALs in our case).

The final model is built using the parameter combination with the highest average accuracy (Figure 3.3). The locked away part of the dataset - the test set - is used to qualify the performance of this model (Kuhn & Johnson 2013). As we are only

Table 3.2: Confusion matrix for landmark identification.

Model	LM	NAL
LM	Identified LM	Unidentified NAL
NAL	Unidentified LM	Identified NAL

interested in landmark identification, the test set includes only landmarks. We use the final model to identify LMs of the test set and count *Identified LMs* and *Unidentified LMs*. A performance measure considering only LMs is the *recall* (Buckland & Gey 1994):

$$Recall = \frac{Identified\ LMs}{Identified\ LMs + Unidentified\ LMs} \tag{3.6}$$

The *recall* does not consider NALs but gives us a hint about the proportion of LMs that has been identified by the model. The trained model is further investigated using sensitivity analysis (Section 3.3) and the results are further evaluated with a McNemar's test (Section 3.4).

3.2.2 'Training' of Models based on Theory

Models based on theory differ from machine learning models as the model is based on explicitly predefined models. This means they do not learn their behaviour from data. However, inspired by the traditional machine learning approach we divide our dataset in two sets (Figure 3.4). The training set in order to create the model and the test set to qualify the performance of the model (Kuhn & Johnson 2013). We do not need NALs for training and testing for the models based on theory. Therefore, their datasets include only landmarks.

Grid-search with cross-validation is identified as a useful practice to find the optimal model parameters for traditional machine learning approaches (Chicco 2017). Following this practice we split the training set into folds using stratified 10-fold cross-validation to get training and validation folds (Stone 1974, Geisser 1975). Again we use the stratified approach as recommended by Kohavi (1995). To do so, we need *classes* to be able to build stratified folds with equal class distributions. Traditional machine learning approaches usually take their target values as classes. The models based on theory do not have target values. Therefore, we need other *classes* to be able to build disjoint folds with equal class distributions. An important prerequisite for the training's success is the availability of landmarks of different decision points of the training set in the folds. Thus, we use the decision points as classes in stratified cross-validation for the models based on theory to be sure to have folds having the same proportion of landmarks from one decision point.

For the traditional machine learning approach we consider both the training and the validation folds (Section 3.2.1). Since the models based on theory are not learned from data we neglect the training folds and only take the validation folds to get an estimate of built models and their performance (Figure 3.4).

We build different personalised wSms and wPms respectively with different model parameter combinations. As the models based on theory do not consider NALs, we only consider the LMs in our performance measure. Therefore, we calculate the recall (Formula 3.6) instead of the accuracy of the built models and select the built model that achieves the highest average recall (Figure 3.4). As recommended by Hsu et al. (2016) for traditional machine learning approaches, we start for the personalised wSm and the wPm also with a coarse grid-search, and after identifying a good region on this grid, a finer grid-search on that particular region follows.

The personalised dFc does not have model parameters because it is built on decisions and processes. Hence, we vary the flow of the model to 'train' it and calculate the recalls of the validation folds. The flow obtaining the highest average recall is the best personalised dFc.

The training results are models based on theory that identify landmarks based on input data. We use the test set to investigate the performance of the models on new unseen data (Figure 3.4). We count the *Identified LMs* and *Unidentified LMs* and calculate the *recall* (Formula 3.6). The trained models based on theory are further investigated using sensitivity analyses (Section 3.3) and the results are further evaluated with a McNemar's test (Section 3.4).

3.2.3 Division in Training and Test Set

For training and testing of the machine learning models we need two independent datasets (James et al. 2013). It is challenging to estimate the optimal ratio for the division of the initial dataset in training and test set. There is no official rule of thumb on the split ratio for training and test set (Wang et al. 2018). Most of the community uses ratios of 50:50 or 80:20 (Sa et al. 2017). Previous research indicates that the test set ratio is proposed to be inversely proportional to the square root of the number of freely adjustable parameters if this number is greater than one (Guyon 1997, Amari et al. 1997).

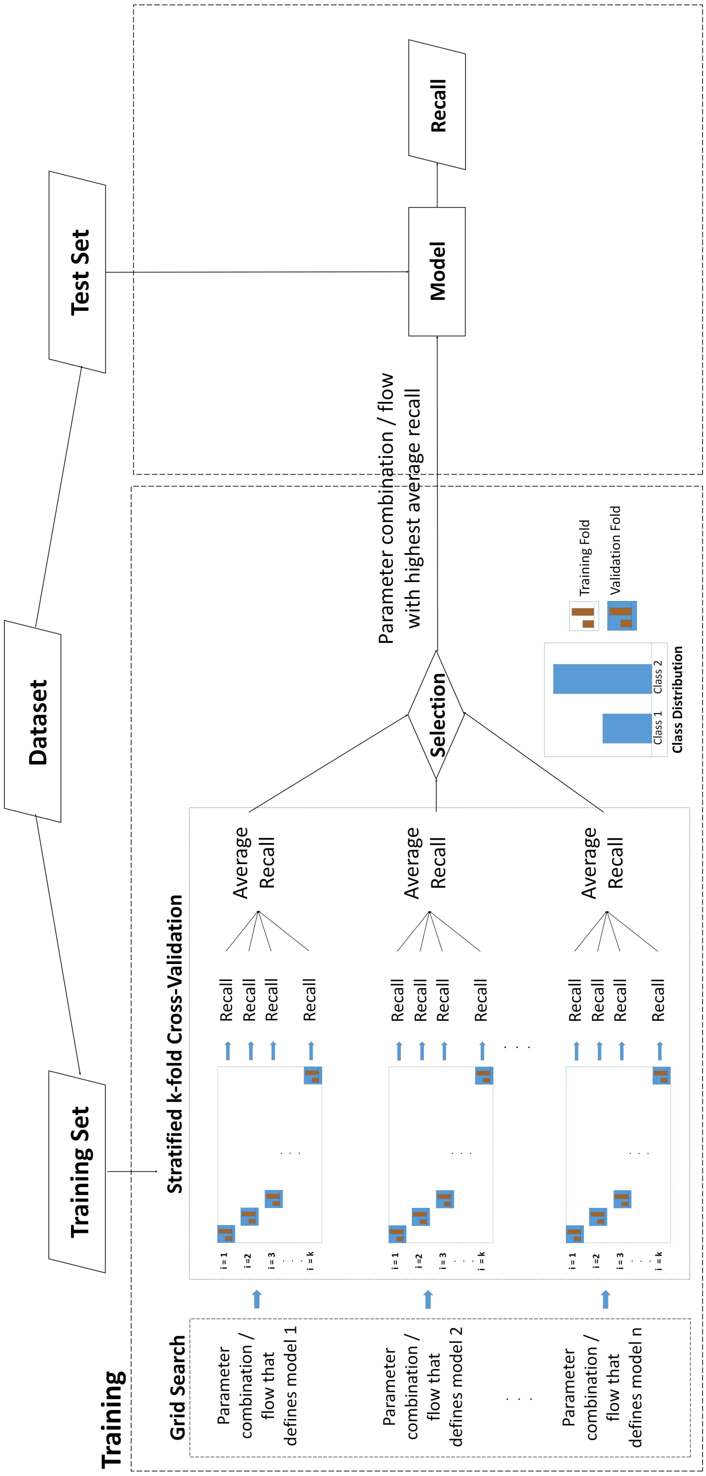


Figure 3.4: 'Training' of models based on theory.

There are several methods to choose independent training and test sets (Bahn & McGill 2013): independent collected data, temporally independent data, and spatially independent data. In case the intended application of the model is to make predictions in new geographic spaces, then spatially independent data should be chosen (Bahn & McGill 2013). We will investigate in Section 5.3 which method and ratio are suitable to split our dataset.

3.3 Sensitivity Analysis of the Models

'The parameter values and assumptions of any model are subject to change and error' (Pannell 1997, p. 139). Sensitivity analysis investigates these changes and errors and their impact on the results of the model (Baird 1989). It investigates how the change of the inputs affects the output of a numerical model (Pianosi et al. 2016). Its importance is widely recognised in several disciplines (Fiacco 1983). Sensitivity analysis follows a simple idea: change the model and observe the results (Pannell 1997). There are many different approaches (for an overview see Pannell (1997)) divided in *local* and *global* sensitivity analysis methods (for an overview see Morio (2011)). For the global analyses all parameters are allowed to vary, whereas the local sensitivity analyses involve variation of only one input parameter at a time which then enables to analyse the effect on the output (Saltelli et al. 2008, Homma & Saltelli 1996). We are interested in identifying the dimensions which actually impact the model results and which do not. Therefore, we perform local sensitivity analysis and vary only one dimension from its minimum value to its maximum value at a time while keeping the values of the other dimensions constant, and then we investigate the outputs. We apply a sensitivity index (SI) to obtain information about the sensitivity of results to different dimensions. There are a number of indices available (Hamby 1994) to measure sensitivity. Comparative assessment of several methods by Hamby (1995) show that the SI proposed by Hoffman & Gardner (1983) performs best. The SI is calculated as follows:

$$SI = \frac{(D_{max} - D_{min})}{D_{max}} \tag{3.7}$$

D_{max}: resulting output value when the dimension is set to its maximum
D_{min}: resulting output value when the dimension is set to its minimum

The SI gives information about the magnitude of differences and the direction in which the model results changes (Jonietz 2016). We determine in Section 7.1 whether

Table 3.3: McNemar's contingency table.

		Model M_2	
		Identified	Unidentified
Model M_1	Identified	N_{++}	N_{+-}
	Unidentified	N_{-+}	N_{--}

different values of the dimensions affect the outputs of the personalised landmark identification models.

3.4 Comparison of Model Results

The choice of the right statistical test for the comparison of model results is a challenging problem (Brownlee 2019). Dietterich (1998) recommends to use McNemar's test in cases where the models compared are only evaluated on one test set instead of repeated evaluations. A contingency table summarises the results for any two models M_1 and M_2 (Everitt 1992) (Table 3.3).

The total number of landmarks in the test set results from

$$n_{Landmarks} = N_{--} + N_{+-} + N_{-+} + N_{++}. \tag{3.8}$$

N_{--}: number of unidentified LMs by both models (M_1 and M_2).
N_{+-}: number of identified LMs by M_1 but unidentified by M_2.
N_{-+}: number of unidentified LMs by M_1 but identified by M_2
N_{++}: number of identified LMs by both models (M_1 and M_2)

The null hypothesis of McNemar's test claims that the two models have the same performance meaning that the number of unidentified landmarks by M_1 but correctly identified by M_2 equals the number of unidentified landmarks by M_2 but correctly identified by M_1 (Dietterich 1998). The null hypothesis of McNemar's test is given by $H_0 : N_{-+} = N_{+-}$ and the alternative hypothesis is $H_A : N_{-+} \neq N_{+-}$ (Kim & Lee 2017).

McNemar's test statistic follows a chi-square distribution with one degree of freedom and is calculated as follows (Kim & Lee 2017):

$$MN^2 = \frac{(N_{+-} - N_{-+})}{N_{+-} + N_{-+}} \tag{3.9}$$

Table 3.4: Example McNemar's contingency table.

		Model M_2	
		Identified	Unidentified
Model M_1	Identified	9	9
	Unidentified	7	17

In McNemar's test discrete values are taken into account. Since the chi-square distribution is continuous, there is an approximation error. To reduce the error, Edwards (1948) proposes a continuity correction. This results in the following test statistic:

$$MN^2 = \frac{(|N_{+-} - N_{-+}| - 1)^2}{N_{+-} + N_{-+}} \tag{3.10}$$

If the null hypothesis is correct, the probability that the test statistic is greater than $\mathcal{X}^2_{1,0.95} = 3.841459$ is less than 0.05 (Dietterich 1998). The p-value calculated by the test can be interpreted as follows with regard to a given significance level α (Brownlee 2019):

- p $> \alpha$: fail to reject H_0, no difference in the performance of the models.
- p $\leq \alpha$: reject H_0, significant difference in the performance of the models.

Table 3.4 shows an example. Amongst these data 17 landmarks being unidentified by both models (M_1 and M_2), seven are unidentified with M_1 but correctly identified with M_2, nine are unidentified by M_2 but correctly identified by M_1 and nine are correctly identified by both models.

A McNemar's test of these data gives the following result:

$$MN^2 = \frac{(|9 - 7| - 1)^2}{9 + 7} = 0.063 \tag{3.11}$$

This has an associated two-tailed p-value of 0.8026. Thus, p $> \alpha$ in case we apply $\alpha = 0.05$. In our example we fail to reject H_0 and cannot detect any statistically significant difference in the performance of the models. We use the McNemar's test for comparing the conventional and personal models among themselves (Section 5.4.5 and Section 6.2). Furthermore, we will compare the results of the conventional models with the results of the personalised models to test our hypothesis (Section 7.2).

3.5 Study Setup

In this chapter we investigated mathematical models and analysis methods as a basis for further investigations. We introduced three models based on theory and one machine learning model (Section 3.1). In the next chapter we will build conventional and personalised landmark identification models (Section 4.3). We will build a conventional weighted sum model (CwSm), a conventional weighted product model (CwPm), a conventional decision flow chart (CdFc), and a conventional decision tree model (CdTm) based on landmark dimensions. In addition, we will build a personalised weighted sum model (PwSm), a personalised weighted product model (PwPm), a personalised decision flow chart (PdFc), and a personalised decision tree model (PdTm) including personal dimensions but based on the conventional models. We intend to compare landmarks identified with the conventional and the personalised models to those selected by survey participants. In the framework of a survey we will collect data (Section 5.1.2). We will divide the resulting dataset into a training and a test set using the methods provided in Section 3.2.3. Then, we will train the machine learning models both, conventional and personalised, on the training set following the traditional approach for the machine learning model (Section 3.2.1).

The conventional models based on theory have no unknown model parameters, whereas the model parameters of the personalised models that are also based on theory need to be identified. The task at hand is to identify the weights of both, PwSm and PwPm, as well as an optimal flow of the PdFc. Inspired by the traditional machine learning approach, we use the 'training' method provided in Section 3.2.2 for the models based on theory. The training results are models that identify landmarks based on input data. Subsequently, we will feed each model with the test set to identify landmarks and calculate their recall (Formula 3.6).

Then, we will analyse the models and their results. We will perform a sensitivity analysis using the methods proposed in Section 3.3 in order to investigate whether changes in the inputs of the dimensions affect the outputs of the models. Afterwards, we will compare the landmarks collected by the survey with the identified landmarks of the models - conventional as well as personalised - and determine whether the collected landmarks are identified correctly or remain unidentified. To detect whether there are any statistically significant differences in the performances of our models we will apply the McNemar's test (Section 3.4).

Chapter 4

Landmark Identification Models

Chapter 3 introduced mathematical models as basis for conventional and personalised landmark identification. Landmarks may have different *dimensions* explaining their *landmarkness*. The first Section 4.1 of this chapter discusses the dimensions considered in the models and identifies attributes for them. Based on the property of salience that turns a conventional geographic object into a landmark, we investigate *salience measures* for all attributes in Section 4.2. The final section presents the models for the identification of landmarks both conventional and personalised (Section 4.3).

4.1 Dimensions of Landmark Identification Models

This section identifies dimensions as basis for landmark identification models. There are dimensions that are dependent on the landmark itself and personal dimensions dependent on the individual traveller. The conventional models consider only landmark dimensions, whereas the personalised models consider both landmark and personal dimensions. This section investigates and discusses the corresponding attributes of the dimensions.

4.1.1 Landmark Dimensions

This thesis builds on the definitions of Sorrows & Hirtle (1999) and Raubal & Winter (2002) for landmark dimensions. The models define the landmark dimensions visual, semantic, and structural dimension. Additionally, we add a dimension to consider the topic of interest.

Figure 4.1: Examples of objects with irregular surface structures.

Visual Dimension

There are four attributes of the visual dimension: *surface structure*, *surface area*, *height*, and *colour*. Raubal & Winter (2002) refer to the *façade* of a building. Because this thesis considers also other urban structures, the attributes relate to the visible *surface* of an object.

Surface structure Objects with irregularly shaped façades or surfaces are easier to recognise than objects with even surfaces. People tend to notice buildings that have salient façades in comparison to the façades of neighbouring buildings (Nothegger et al. 2004). Buildings are visually salient if they show e.g. bay windows, balconies, or outstanding façades. Surfaces of other objects are irregular if they are not uniformly shaped (e.g. a water wheel with its blades or an advertisement pillar with different colours (Figure 4.1)).

Surface area Another attribute that classifies an object as salient, is one with a surface different from all the others. Already the participants in Lynch's (1960) study about the image of a city called 'varied roof tops' (p. 162) as an important aspect. A building with a tent roof in a neighbourhood where saddled roofs are dominant is outstanding (Figure 4.2a). Other objects such as stationary bollards with round surface areas (Figure 4.2b) or a street light with a peaked roof (Figure 4.2c) might be considered as outstanding.

Height A different height from all the other surrounding objects can give an object a salient appearance. For example, television towers, hilltops, and city skylines might be valuable global landmarks (Steck & Mallot 2000). Vice versa: small objects (monuments, garbage bins, or park benches) might be outstanding because of their height. A variation in height of local objects sets up a contrast with nearby elements (Lynch 1960).

(a) Tent. (b) Round. (c) Peaked.

Figure 4.2: Examples of objects with outstanding surface areas.

Colour Colour is another attribute of the visual dimension. An object can stand out because of its colour from surrounding objects (Raubal & Winter 2002). Colour is a cure for structuring and identifying the environment (Lynch 1960). For example, a blue house in a street with grey houses would attract attention from a traveller. A red telephone box in an otherwise grey environment might be visually attractive because of its colour.

Semantic Dimension

In this thesis the use of the notion of semantic attraction of an object is the same as Raubal & Winter (2002) and similar to that of cognitive attraction (Sorrows & Hirtle 1999). The models consider *cultural*, *historical importance*, and *explicit marks* as attributes of the semantic dimension.

Cultural importance Landmarks are defined by a combination of attributes including cultural importance (Sadalla et al. 1980). An object is culturally important if it promotes culture or arts or is a place of leisure or entertainment. This includes buildings that accommodate sport centres, public swimming pools, cinemas, or museums, but also places such as parks, entertainment areas, or marketplaces.

Historical importance Semantic attraction of an object might result from historical importance (Sorrows & Hirtle 1999). Buildings with historical importance often stand out because of their architecture (e.g. in an urban environment city walls or old historic buildings). In addition, structures that are different from buildings such as monuments with a historic meaning or historic places have a certain importance.

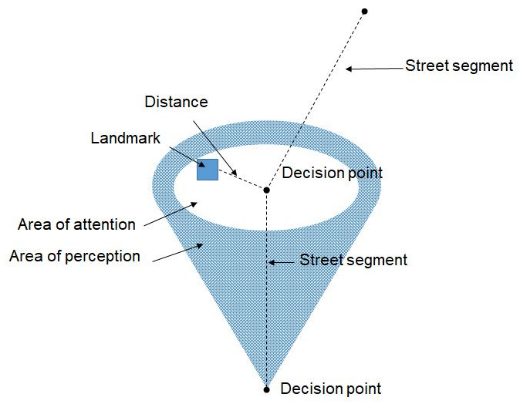


Figure 4.3: Attributes of the structural dimension (modified from Maass (1996)).

Explicit marks Objects with explicit marks are of great value to a traveller because they are easy to identify. An explicit mark on an object specifies its semantics to the traveller (Raubal & Winter 2002). For example, when a building is marked as *Museum*, then its use is immediately apparent. Street signs or monuments with explicit marks might be valuable navigation aids as well. Explicit marks are an additional information of an object, which cannot be identified solely by its visual attributes.

Structural Dimension

Objects are structurally salient as soon as they have prominent spatial locations within an environment (Sorrows & Hirtle 1999) or if they are highly accessible. Attributes of the structural dimension are *location at a decision point* and *distance to the decision point*.

Location at a decision point Decision points are of particular importance because they are mostly linked to actions, such as 'turn left' or 'go straight' (Montello 1998). In order to take account of the fact that objects at decision points are more valuable for route directions, objects are salient if they are located at decision points.

Distance to the decision point Objects close to a decision point are useful for navigation purposes (Waller et al. 2000). A traveller normally restricts his attention to an area of perception around a decision point (Maass 1996). Within this area of perception the traveller focuses on a spatial area of attention (Figure 4.3). Objects near a decision point and within the area of attention are preferred as a navigation aid.

Landmark Interest Dimension

An object might be outstanding because of the topic of interest it belongs to. There are a number of topics of interest that may be attached to urban objects, such as gastronomy or shopping but also historical or cultural interest. A building hosting a restaurant may attract attention for people who like to go out and visit a restaurant. Other urban objects might be attractive because of their history, such as city walls or monuments.

In Section 5.1.1 we analyse objects of Augsburg's innercity because we need to identify possible topics of interest. Two of them are the topics of interest *cultural interest* and *historical interest* (Table 5.1). They overlap with the attributes *cultural importance* and *historical importance* of the semantic dimension. However, the information on the topic of interest needs to be explicitly available for the assessment of the personal dimension *personal interests* in the PdFc (Section 4.3.6). It investigates whether a traveller has some *personal interest* in a topic. In this case the model selects objects that are part of this particular topic. To enable the PdFc to do so, the objects must be assigned to those topics of interest they belong to. The *landmark interest dimension* provides the information to which topics of interest an object belongs to.

4.1.2 Personal Dimensions

This section deals with the definition of personal dimensions of objects for the inclusion within a personalised landmark identification model. In general, five dimensions are important when viewing a person as an individual (Brusilovsky & Millán 2007):

1. Personal knowledge,
2. Personal interests,
3. Personal goals,
4. Personal background, and
5. Individual traits.

There might be more other not yet identified personal dimensions. However, for this first approach on modelling personalised landmarks we base ourselves on the dimensions provided by Brusilovsky & Millán (2007).

Probably the most important dimension to consider for personalised landmark identification is personal knowledge. In this thesis personal knowledge refers to the *prior spatial knowledge* of the traveller. A highly personalised landmark such as *my working place* or *the home of a friend* is a good landmark only if the traveller has spatial knowledge of the environment. *Personal interests* constitutes an important dimension, because a travellers level of interest enhances memory for some information (McGillivray et al. 2015). The traveller's *personal goal* represents the immediate purpose for a traveller's work with a personalised application (Brusilovsky & Millán 2007). It is the most variable dimension of the above-mentioned ones and has an impact on the amount of required route directions and on the distribution of landmarks. The *personal background* of the traveller is a common name for a number of attributes related to the travellers previous experiences outside the landmark or navigation domain (Brusilovsky & Millán 2007). The personal background influences the way objects are recognised and perceived. *Individual traits* subsume the features of the traveller that together define a traveller as an individual and might influence how objects are perceived. The following sections discuss the personal dimensions and their attributes in detail.

Prior Spatial Knowledge

Probably the most important dimension to consider for the provision of personalised landmarks is the *prior spatial knowledge* of a traveller. We use discrete qualitative categories based on the *dominant* framework proposed by Siegel & White (1975) (Section 2.1.1). This thesis proposes landmark, route, and survey knowledge as attributes of the dimension *prior spatial knowledge*. In addition, it adds a fourth attribute - *no knowledge*, for those areas where the traveller has never been before.

Landmark knowledge Travellers notice various objects and encode images of the environment while first encountering an unfamiliar area (Thorndyke 1980). Von Stülpnagel & Steffens (2013) show that self-contained movement through an environment leads to the encoding of landmark knowledge. Then, people are able to recall the objects they have seen and to remember e.g. names of certain buildings and locations. These objects are the first spatial cues that are available in no particular order on a cognitive map (Couclelis et al. 1987). Travellers with only landmark knowledge are not familiar with routes and not able to find short cuts and detours although they are not completely unfamiliar with the environment to navigate.

Route knowledge Previous experiences with a route through an environment lead to changes in potential wayfinding effectiveness (Allen & Kirasic 2003). Route knowledge enables to navigate from a starting point to a destination without any aids (e.g. maps, navigation systems, route directions). The availability of knowledge of a sequence of landmarks along a route and the knowledge of how to get from one landmark to another influences the granularity of route directions (Tenbrink & Winter 2009). There are many reasons to remember objects previously used in route directions (Winter 2003). Objects next to an already navigated route are structurally salient and their location along the route is cognitively easy to conceptualise (Klippel & Winter 2005).

Survey knowledge Survey knowledge is usually generated from route knowledge through integration of the routes into a cognitive map (Tolman 1948, Downs & Stea 1974, O'keefe & Nadel 1978). Survey knowledge implies that the traveller is familiar with a certain environment. Quesnot & Roche (2015) show that travellers familiar with a specific environment prefer objects with semantics as landmarks. Such objects have personal meanings solely because of their semantics, e.g. *the place where I once lived* or *my doctor* (Richter & Winter 2014) - even a bright-coloured door might be a landmark if it is for example your own (Lynch 1960). The higher the degree of familiarity the higher is the possible degree of personalisation of a landmark. For example, the house of a friend in a building ensemble, although structurally and visually identical to the other buildings, may become a landmark. These personal semantic or cognitive landmarks might be missed by travellers unfamiliar with the environment, unless there are some explicit marks (Sorrows & Hirtle 1999).

No knowledge The fourth attribute of the personal dimension *prior spatial knowledge* is no knowledge. In case travellers have never been to the environment to navigate before and have never seen a map or photos, then we assume that they have no *prior spatial knowledge* at all. Quesnot & Roche (2015) show that people not familiar with an environment prefer landmarks because of their visual or structural salience. For these travellers highly visible landmarks located at strategic decision points of the route should be provided.

Personal Interests

Undoubtedly, preferences for certain objects and for activities involving the objects exists (Fink 1991). Travellers turn their attention towards certain objects in their environment to keep these objects within their field of perception. Travellers must look around in order to perceive things, especially when navigating with landmarks. They walk through the streets keeping their eyes open for the next landmark. But looking around is not enough to perceive objects (Rensink et al. 1997). A traveller whose mind wanders during walking may often miss out on important things, even when these are highly salient (Simons 2000). People are easy to distract and as a consequence they may miss important objects, such as landmarks (Arthur & Passini 1992). Rensink et al. (1997) argue that the key factor for perceiving things is attention, which is dependent on the degree of interest. Without attention many people have no awareness at all of some objects along a route and are *inattentionally blind* (Rock et al. 1992, Mack et al. 1998). Banerjee et al. (2015) confirm that the observer's level of interest in an object influences the voluntary focus of attention on environmental inputs. They show that participants perform better in a visuospatial task of spatial target detection of high interest items. *Personal interests* may result in selective attention which is related to the locus of eye fixations during navigation. This, according to Viaene et al. (2016), may point to recognition or use of a landmark for wayfinding. In addition, studies show that people show an improved memory for information that they are curious about (Gruber et al. 2014, McGillivray et al. 2015).

Interests are conceptualised from either a situational perspective or a personal perspective (Schraw & Lehman 2001, Hidi & Renninger 2006). Some authors differentiate between preferences and interests (Weiβenberg et al. 2006). Preferences or situational interests are caused by certain conditions and/or concrete features of the environment (Renniger & Su 2012). They are dependent on the situation of the traveller and external factors. There are simple and complex preferences (Weiβenberg et al. 2004). For example, consider travellers not much interested in historical monuments. During their holidays they may prefer these monuments to get to know the culture; this is referred to as *complex* preference. On the other hand, there are *personal interests* (also referred to only as interests (Weiβenberg et al. 2004)), which reflect personality-specific orientation. *Personal interests* are static and application specific parameters and are defined before using an application (Weiβenberg et al. 2004). *Personal interests* provide important categories for action goals in a situation when travellers are free to do as they please (Krapp et al. 2017). There are many different possible interests for a traveller in an urban environment. For example, a traveller who loves to go to bars and restaurants, but is bored by art and culture, will obviously be more attentive to gastronomy-related objects than urban features such as statues or monuments.

Personal Goals

Objects are classified as landmarks as soon as they are helpful aids to achieve a goal. The salience of an object is not dependent on the traveller's goal because it depends much more on the object's attributes and on personal dimensions (e.g. *prior spatial knowledge* or *personal interests*). It is, however, important to know the traveller's goals, because it makes a difference in the number and the distribution of landmarks. In human wayfinding three goals are distinguished: travel with the goal of reaching a familiar or a novel destination and exploratory travel (Section 2.1.1). Depending on the particular situation travellers pursue different goals.

Known goal When travellers are navigating to a known goal their focus is primarily on reaching a particular location. This may be a specific spatial goal (e.g. *the house of a friend*). In this case, travellers need no landmarks around the destination, because they are already familiar with it. The distribution of landmarks along the rest of the route depends on the familiarity of the traveller with the route.

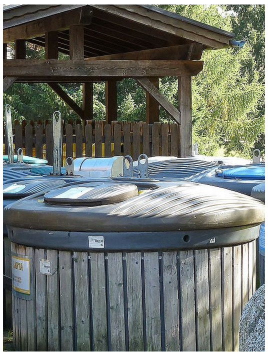
New goal Travelling with the goal to reach a novel destination is mostly carried out with different kinds of aids (Section 2.1.1). There is a need for more landmarks around the destination when it is unknown. Michon & Denis (2001) show that the frequency with which landmarks are mentioned in route directions increases in the vicinity of the destination. At points where a change in direction is required or along long route segments confirmatory landmarks should be provided. We assume that if the goal is unknown, the route or at least parts of the route are unknown as well.

Exploratory travel The situation during exploratory travel is different to travelling to a known or unknown goal. In this case travellers may be interested in extra landmarks along the route. Additionally, informative landmarks are helpful to get to know the environment.

Personal Background

Personal background is a common name for attributes describing a traveller's experience outside of a specific application (Brusilovsky & Millán 2007) - in this case navigation and wayfinding. The personal background is mainly described by demographic data - *objective facts* (Kobsa et al. 2001) - and gives information about the personal characteristics of a traveller. Data describing the personal background may include record, geographic, psycho-graphic, or customer qualifying data as well as data

(a) The Netherlands.



(b) Spain.

(c) Germany.

(d) Germany.

(e) United Kingdom.

Figure 4.4: Examples of objects with different sizes, shapes, or colours.

describing the travellers characteristics (Kobsa et al. 2001). Important demographic data are geographic data (country of residence, cultural background) and data about the traveller's characteristics (education, gender, age).

Country of residence The first geographically-related attribute is the country of residence. It is an important attribute, because travellers not living within the country of the environment to navigate, may be used to environments and objects shaped differently (Kattenbeck 2016). There are quite a number of objects which have different sizes, shapes, or colours in different countries (e.g. compared to Germany, telephone boxes in the United Kingdom have a different colour and shape and glass containers have a different size and shape in the Netherlands and Spain (Figure 4.4)).

Cultural background The second geographically-related attribute is the cultural background of the traveller. Here, the same applies as for the country of residence: travellers, who did not grow up within the environment to navigate may be used to completely different objects and shapes. Consider travellers who grew up in a small rural village, they have a different background compared to travellers who grew up in the middle of a large modern city.

Education There are attributes of the personal background important for the identification of personalised landmarks concerning the traveller's characteristics. One of them is the education of the traveller. Berry & de Rosis (1991) reveal that a user's knowledge in a domain varies considerably according to their background and job. Concerning navigation and wayfinding, the education of a traveller may influence the way visual and structural dimensions are perceived (Kattenbeck 2016). Consider e.g. sculptors that have a perspective on statues or art work or surveyors who take special note of measuring points or benchmarks whereas others do not even notice these spatial objects.

Age A further attribute concerning a traveller's characteristics is age. The age of travellers is found to be an important attribute in spatial cognition because of strong differences in orientation abilities (Jansen-Osmann et al. 2007) and route memorisation (Wang et al. 2019). Jansen-Osmann & Wiedenbauer (2004) show that younger people rely more on the presence of landmarks than adults. Goodman et al. (2005) show that a pedestrian wayfinding application including landmarks is particularly useful for older people and indicate a need for personalisation for elderly people. Age may have a particular impact on the structural salience of objects (Kattenbeck 2016).

Gender The third attribute of the traveller's characteristics is gender. There are known differences regarding spatial cognition between women and men (Coluccia & Louse 2004, Wang et al. 2019) and use of landmarks (Ward et al. 1986). Wang et al. (2019) state that males pay less visual attention to landmarks than females. Other studies report differences between women and men in the importance of structural salience (Quesnot & Roche 2015).

Individual Traits

Individual traits deal with the attributes of travellers that define them as an individual (Brusilovsky & Millán 2007). Examples are personality traits (e.g. introvert/extrovert), cognitive styles (holist/serialist), cognitive factors (e.g. working memory capacity), and learning style (Brusilovsky & Millán 2007). Individual traits are stable parameters of a traveller that either do not change at all or only change over a long period of time. These parameters might be identified with specially designed psychological tests (Brusilovsky & Millán 2007). Existing work on modelling individual traits for personalisation mostly deals with cognitive styles and learning strategies (Riding & Rayner 1998). Goren-Bar et al. (2006) investigate personality traits together with adaptivity, which is a 'technological approach whereby systems monitor and manipulate personal needs and interests' [p. 32]. They show that personality traits relating to the notion of control have a selective effect on adaptivity acceptance. They outline that any evaluation of a mobile application might be biased unless the personality of the users is taken into account.

4.2 Salience Measures for the Dimensions

The goal of the landmark identification models is either to determine an overall salience measure of landmarks or to classify objects in *landmark* (LM) and *not a landmark* (NAL). These calculations are based on salience measures for the visual (vis), the semantic (sem), and the structural (str) dimension. In this section we investigate salience measures for the attributes of the landmark dimensions. Furthermore, we investigate measures to consider the salience of personal dimensions.

4.2.1 Landmark Dimensions

This thesis assigns landmark salience values in percent to the object as soon as an attribute value is different or differs from the attribute values of the surrounding objects. In case all attribute values of a dimension are salient, the object gets a 100% salience for this dimension. Consider e.g. an object meeting all the requirements of the visual dimension, then it is awarded 100% visual salience. An object that is for example only visually attractive because of its surface area and structure only gets a 50% salience. An object must fulfil specific conditions to be considered salient (Table 4.1). Salience is based on threshold values from Raubal & Winter (2002) and Nuhn et al. (2012).

Visual Dimension

For the attributes of the visual dimension, surface structure, surface area, height, and colour threshold values are defined indicating when their values differ significantly from the values of the surrounding objects in a local neighbourhood. The local neighbourhood may be a buffer of a specific size. This thesis follows Raubal & Winter (2002) and assumes that each of the attributes of the visual dimension have the same effect on the overall salience of an object. We assign a salience value of 25% in case the attribute is salient (Table 4.1, column *Salience (Attribute)*). Zero percent means that the attribute is not salient.

Table 4.1: Rules for the computation of landmark salience.

Dimension	Attribute	Salient	Salience (Attribute)	Salience (Dimension)
Visual	Surface Structure V_s	If *True*	$s_{V_\mathrm{s}} \in \{0, 25\}$	$s_\mathrm{vis}[\%] = s_{V_\mathrm{s}} + s_{V_\mathrm{a}} + s_{V_\mathrm{h}} + s_{V_\mathrm{c}}$
	Surface Area V_a	See text below	$s_{V_\mathrm{a}} \in \{0, 25\}$	
	Height V_h		$s_{V_\mathrm{h}} \in \{0, 25\}$	
	Colour V_c		$s_{V_\mathrm{c}} \in \{0, 25\}$	
Semantic	Cultural importance S_c	If *True*	$s_{S_\mathrm{c}} \in \{0, 25\}$	$s_\mathrm{sem}[\%] = s_{S_\mathrm{c}} + s_{S_\mathrm{h}} + s_{S_\mathrm{e}}$
	Historical importance S_h		$s_{S_\mathrm{h}} \in \{0, 25\}$	
	Explicit marks S_e		$s_{S_\mathrm{e}} \in \{0, 50\}$	
Structural	Location at a Decision Point St_l	If *True*	$s_{St_\mathrm{l}} \in \{0, 50\}$	$s_\mathrm{str}[\%] = s_{St_\mathrm{l}} + s_{St_\mathrm{d}}$
	Distance to the Decision Point St_d	If $St_\mathrm{d} = \min(St_\mathrm{d1},...St_\mathrm{di})$	$s_{St_\mathrm{d}} \in \{0, 50\}$	
Interest	Belonging to a topic of interest I_LM	If *True*	$s_\mathrm{iLM} \in \{0, 1\}$	s_iLM

Surface structure A building with an outstanding façade or an object with an irregular surface gets the Boolean value *True* - and *False* otherwise. The surface structure is salient as soon as this building or object has an attribute value *True*.

Surface area The value of the attribute surface area is a String describing the kind of surface area (e.g. *tent*, *flat* for buildings or *round*, *peaked* for other objects). The surface area is salient as soon as the String value is different from all the others in a local neighbourhood. That means, e.g. if a building with a tent roof is classified salient because of its surface area, it is the only one with that kind of roof.

Height Each object has its individual height. The attribute value of the attribute height is a number. The assessment whether this attribute value is significantly different from mean characteristics within a local neighbourhood is done by hypothesis testing. The null hypothesis claims that the attribute value of height is not significantly different from the others. In case the null hypothesis is rejected, the object has a significant height.

Colour The attribute value of colour is a String (e.g. red, blue, or yellow). An object is salient because of its colour being different from all the other colours of the objects in a local neighbourhood. This might be for example, a telephone box with the colour red, whilst all the other objects around do not have the colour red.

Semantic Dimension

The salience of the attributes of the semantic attributes - cultural and historical importance as well as explicit marks - is independent of the other objects in the neighbourhood. They are measured with Boolean values and considered as salient if they are *True*. The attributes get different salience values (see below). The semantic salience is zero if there are respectively neither cultural nor historical importance nor explicit marks.

Cultural importance Cultural importance receives a salience of 25% (Table 4.1, column *Salience (Attribute)*) if the object has a cultural value. An object that does not promote culture gets the salience value zero.

Historical importance This thesis assigns a 25% salience to historical importance if the objects are meaningful in history. If the value for this attribute is *False*, the objects get no salience for historical importance.

Explicit marks We assume that the availability of explicit marks is of a higher value than cultural or historical importance. Therefore, explicit marks get a percentage salience value of 50% as soon as there is an explicit mark available.

Structural Dimension

The attributes of the structural dimension must meet certain conditions to be salient. Similar to the case of visual attributes of the landmark dimension this thesis assumes that each of the attributes has the same effect on the overall salience of the dimension and assigns a salience value of 50% (Table 4.1, column *Salience (Attribute)*). If the attribute of an object is not salient, it gets zero percent. The following paragraphs explain the conditions that must be met.

Location at a decision point In case an object is located at a potential decision point, it gets the Boolean value *True* for that attribute. More than one object at a decision point can get *True* for that attribute, because there normally is more than one object located at a decision point. An object located at street segments gets the Boolean value *False* for that attribute. Since we focus in this thesis on landmarks at decision points, each object is salient and gets the Boolean value *True* for that attribute.

Distance to the decision point The distance to the decision point is stored as a number. The object with the smallest distance to the decision point gets a percentage of a salience value. The other objects get a salience of zero for the attribute distance to the decision point.

Landmark Interest Dimension

Some objects belong to a number of topics of interest. Consider a restaurant which belongs to the topic of interest *gastronomy*. In case the buildings architecture is outstanding, it might belong to the the topic of interest *architecture*. As soon as an object belongs to a topic of interest, it gets a landmark interest (iLM) salience value of $s_{iLM} = 1$ for that particular interest. Zero means that the object is not interesting for that particular topic of interest.

Table 4.2: Stages of prior spatial knowledge (PspK).

s_{PspK}	Been before at the street intersection?	Knowledge	The traveller...
1	Yes	Survey	... has been in the area before and knows short-cuts and detours.
2		Route	... has been in the area before and knows some routes through the area.
3		Landmark	... only knows some important points in the area.
4	No	Survey	... has been in the area before and knows short-cuts and detours.
5		Route	... has been in the area before and knows some routes through the area.
6		Landmark	...only knows some important points in the area.
7		No	... has never been in the area before.

4.2.2 Personal Dimensions

This thesis is a first approach on modelling personalised landmarks. Following (Klippel et al. 2009) who identified 'user's familiarity with an environment, as well as personal styles' (p. 231) as important aspects of cognitively ergonomic route directions, we focus on *prior spatial knowledge* and *personal interests* in this thesis. We start with these two personal dimensions and concentrate on how they might be incorporated in personalised landmark identification models. We discuss reasons why the other dimensions are treated elsewhere below.

Prior Spatial Knowledge

In Section 4.1.2 we identify stages of *prior spatial knowledge (PspK)*, namely landmark, route, survey, and no knowledge. In addition to these stages we differentiate if the traveller has been before at the investigated decision point or not. We differentiate seven stages of *prior spatial knowledge* (Table 4.2). *Prior spatial knowledge* is an aspect that influences the other dimensions and their attributes. For that reason their salience is not expressed by percentage values but by numbers ($s_{PspK} \in \{1, ..., 7\}$). These numbers are either transferred to weights in the PwSm (Section 4.3.2) and the PwPm (Section 4.3.4) or directly used in the PdFc (Section 4.3.6) and the PdTm (Section 4.3.7).

Personal Interests

We assume that the travellers interests in topics influence their landmark selections. We focus solely on *personal interests (pInt)* dealing with person-specific orientation in general. Situational interest is treated elsewhere (Section 4.1.2). The interest in different topics varies considerably between different travellers. Rating scales are one way to measure *personal interests* which estimate the travellers interest in a topic by a single value on a specific scale. A travellers interest in a topic might range from one (*no* interest), two (*low*), three (*medium*), four (*high*), to five (*very high* interest). This results in a *personal interests* salience $s_{pInt} \in \{1, ..., 5\}$. Another approach consists of not considering scales of interest, but only *interested* or *not interested.* To transfer the five point interest scale to the two point scale we consider $s_{pInt} \in \{1, 2, 3\}$ as *not interested* and $s_{pInt} \in \{4, 5\}$ as *interested.* Then the interest salience is $s_{pInt} \in \{0, 1\}$. The salience is - as was the case with *prior spatial knowledge* - transferred to weights or directly considered in the personalised landmark identification models.

Personal goals

Wayfinding goals have an impact on the number and the distribution of landmarks along a route. The salience of an object is not dependent on the traveller's personal goals for wayfinding. That is why we do not further discuss personal goals explicitly but make references at some points where personal goals might influence the results.

Personal background

We do not discuss personal background further at this point. The incorporation of this dimension would require deeper analysis of its influence on the overall salience of objects. This would involve user studies with psychological tests regarding gender and age and a study on how objects are perceived in different countries from people with different cultural background. This is beyond the scope of this thesis, which gives a first approach on modelling personalised landmarks.

Individual Traits

Unlike the other dimensions individual traits can only be determined through especially designed psychological tests. Such tests would involve sound psychological knowledge, which can only be contributed by experts. For this reason, this thesis does not incorporate individual traits in the models.

4.2.3 Salience Vector

After determining the salience measures for the individual dimensions we model the salience of every object o_i^j from a set O^j of objects for each traveller as a salience vector of dimensions:

$$\vec{s}_{ov} = \begin{pmatrix} s_{vis} \\ s_{sem} \\ s_{str} \\ s_{iLM} \\ s_{PspK} \\ s_{pInt} \end{pmatrix}$$

We assume that at every decision point j at least one object o_i^j is available as input for our models. In the next Section 4.3 we investigate models for identifying landmarks for the decision point.

4.3 Overall salience

This section presents the models we use to identify landmarks - conventional as well as personalised models. The P/CwSm (Section 4.3.1 and Section 4.3.2) and the P/CwPm (Section 4.3.3 and Section 4.3.4) calculate an overall salience measure for each object. The P/CdFc (Section 4.3.5 and Section 4.3.6) classify objects as LMs and the P/CdTm (Section 4.3.7) in LMs and NALs.

4.3.1 Conventional Weighted Sum Model (CwSm)

The CwSm is used for landmark identification by Raubal & Winter (2002). They use the well-established visual, semantic, and structural dimensions but include slightly different attributes from those presented here (Section 4.1). Raubal & Winter (2002) determine values for each attribute and investigate whether an attribute value is significantly different from the others in a given neighbourhood. Therefore, authors use hypothesis testing. They set the significance value to 1 in case there are significant differences, i.e. the attribute is salient for a specific object. Otherwise, the significance value is zero. We use the approach from Raubal & Winter (2002) but include our attributes and consider the attribute values as salient as soon as they fulfil the salience conditions in the column *Salient* in Table 4.1. We use the salience defined in column *Salience (Attribute)* in Table 4.1.

Then, Raubal & Winter (2002) group the salience values for visual, semantic, and structural dimensions (Table 4.1, column *Salience (Dimension)*). They determine the total measure of landmark salience for each object by adding up the grouped salience values (Formula 4.1). The landmark with the maximum overall salience is established as a potential landmark and is used to enrich route directions. Raubal & Winter (2002) set the weights to one. They propose to adapt the weights in their total salience measure to the context or individual user preferences but do not discuss this any further.

$$s_{CwSm} = (w_{vis} * s_{vis} + w_{sem} * s_{sem} + w_{str} * s_{str})/100 \tag{4.1}$$

$w_{vis} = w_{sem} = w_{str} = 1$

We use percentage values for s_{vis}, s_{sem}, and s_{str} to determine landmark salience (Table 4.1). The overall salience measure is divided by 100 in Formula 4.1 for the sake of clarity.

4.3.2 Personalised Weighted Sum Model (PwSm)

The PwSm is quite similar to the CwSm (Section 4.3.1). We adapt the weights in the CwSm for the consideration of *personal interests* as well as *prior spatial knowledge* within the PwSm (Formula 4.2). We assign weights according to s_{pInt} and s_{PspK}. The weights cannot be zero because this results in empty terms for the visual, the semantic, and the structural dimension.

$$s_{PwSm} = (w_{vis} * s_{vis} + w_{sem} * s_{sem} + w_{str} * s_{str})/100 \tag{4.2}$$

$w_{vis} = \mathrm{f}(s_{pInt}; s_{PspK})$
$w_{sem} = \mathrm{f}(s_{pInt}; s_{PspK})$
$w_{str} = \mathrm{f}(s_{pInt}; s_{PspK})$

We divide the overall salience measure in Formula 4.2 by 100 for the sake of clarity. The result of the PwSm is a measure of landmark salience for an object. There might be one or more objects with the highest salience measure at a decision point. In Section 6.1.1 we determine weights for the PwSm.

4.3.3 Conventional Weighted Product Model (CwPm)

We are not aware of any existing wPm to identify landmarks. We build the model similar to the CwSm. We determine the values for each attribute of the dimensions. Then we investigate whether an attribute value is salient according to Table 4.1 and salience values are grouped. The CwPms overall measure of landmark salience considering only landmark dimensions is calculated with weights set to one (Formula 4.3). We divide the salience measure by 100 for the sake of clarity.

$$s_{CwPm} = (s_{vis}^{w_{vis}} * s_{sem}^{w_{sem}} * s_{str}^{w_{str}})/100 \tag{4.3}$$

$w_{vis} = w_{sem} = w_{str} = 1$

4.3.4 Personalised Weighted Product Model (PwPm)

We build the PwPm quite similar to the CwPm (Section 4.3.3). The total measure of personalised landmark salience is gained with weights dependent on s_{pInt} and s_{PspK} (Formula 4.4).

$$s_{PwPm} = (s_{vis}^{w_{vis}} * s_{sem}^{w_{sem}} * s_{str}^{w_{str}})/100 \tag{4.4}$$

$w_{vis} = \mathrm{f}(s_{pInt}; s_{PspK})$
$w_{sem} = \mathrm{f}(s_{pInt}; s_{PspK})$
$w_{str} = \mathrm{f}(s_{pInt}; s_{PspK})$

Again, we use percentage values to determine landmark salience and, for the sake of clarity, divide it by 100 . The result of the PwPm is again a landmark salience measure. There might be, as in the case of the PwSm, more than one object with the highest measure at a decision point. Section 6.1.1 investigates the determination of weights for the PwPm.

4.3.5 Conventional Decision Flow Chart (CdFc)

We build a basic CdFc following the steps in Section 3.1.1 and using the symbols provided. Our area of focus is the identification of an object that is suitable as a landmark. We follow the results from our literature research (Section 2.1.2) to identify the steps of the flow in chronological order. The first process investigates the visual

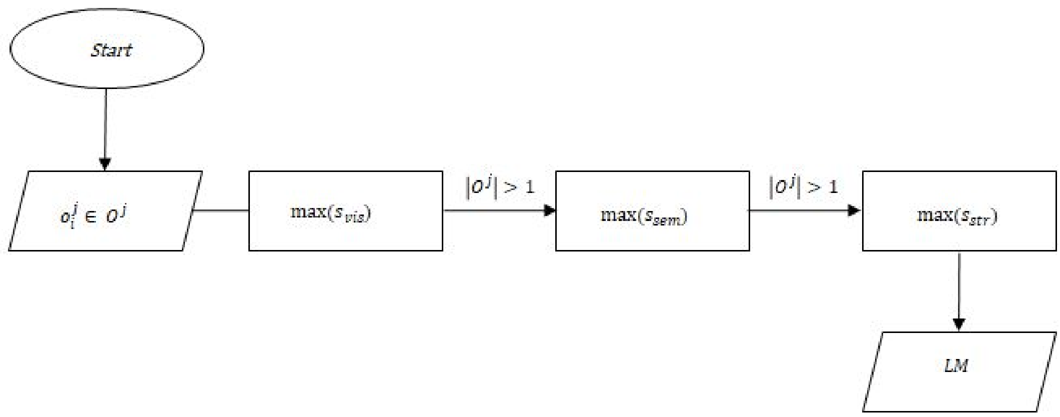


Figure 4.5: Conventional Decision Flow Chart.

salience of an object, followed by a process assessing the semantic salience, and a process determining the object with the maximum structural salience (Figure 4.5).

The input of the CdFc is every object o_i^j from a set O^j of objects at a decision point. It flows from left to right considering the next process provided there is still more than one object available ($|O^j| > 1$). The CdFc directly proceeds to the output (LM) in case there remains only one object as a result of a process. For a better overview these connections are not depicted in Figure 4.5.

4.3.6 Personalised Decision Flow Chart (PdFc)

For the identification of personalised landmarks we build a personalised decision flowchart following the steps in Section 3.1.1 and using the symbols provided. Our area of focus is to identify the most personal object that is suitable as a landmark $o_{i,max(pers)}$. For the identification of the steps in chronological order we follow the results from our literature research (Section 2.3). The most important dimension to consider is the familiarity or the *prior spatial knowledge* of the traveller (Hamburger & Röser (2014), Quesnot & Roche (2015), Caduff & Timpf (2008)). Therefore, the first component of the flowchart deals with the decision about the particular *prior spatial knowledge* at the decision point (Figure 4.6). This *prior spatial knowledge* is reflected in the importance of visual, semantic, and structural salience. We consider landmark dimensions (visual, semantic, and structural) next in the flow, followed by the investigation of *personal interests*.

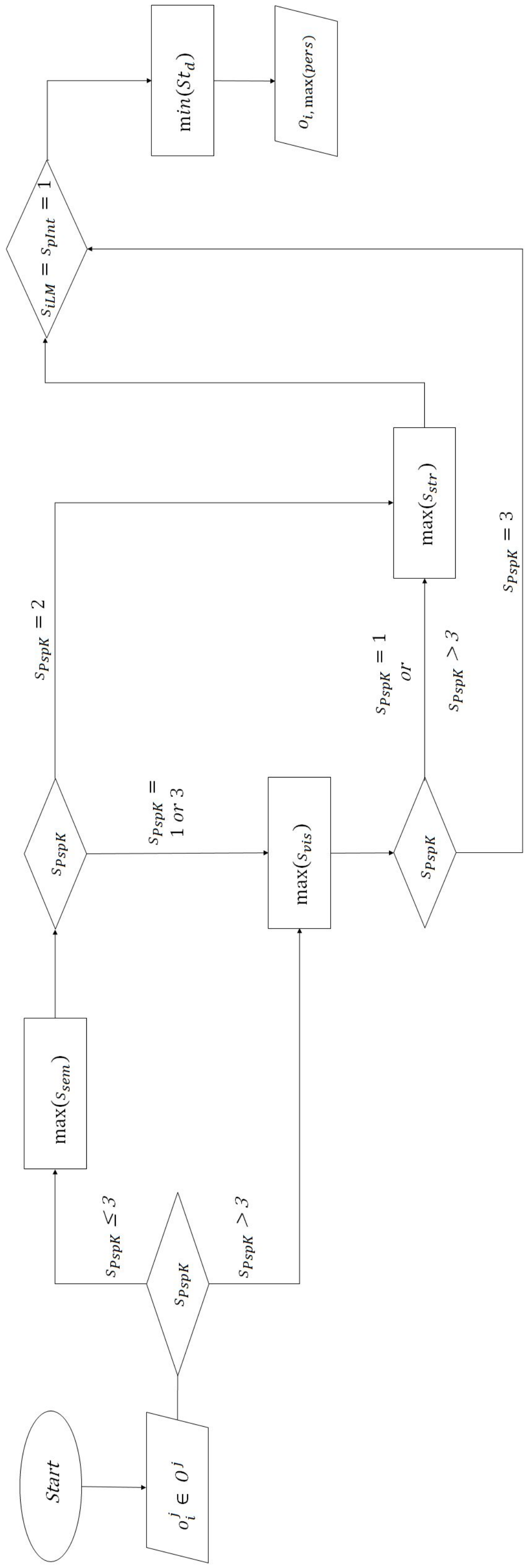


Figure 4.6: Personalised Decision Flow Chart.

The flowchart considers every object o_i^j from a set O^j of objects at a decision point as input (Figure 4.6). It flows from left to right considering the next decision or process provided there is still more than one landmark available ($|O^j| > 1$). In case there remains only one object as a result of a decision or a process, the flowchart directly proceeds to the output ($o_{i,max(pers)}$) and accepts this object as the most personal one. For reasons of simplification these connections are not depicted in Figure 4.6.

The chart starts with the decision if the traveller is familiar with the street intersection. People unfamiliar with an environment use highly visual and structurally attractive landmarks (Quesnot & Roche 2015). Thus, the landmark with the maximum visual salience and then (if there is more than one landmark) the one with the maximum structural salience is determined. If there is more than one landmark passing both processes, the flowchart continues with a decision about the interest salience. The PdFc divides in interested and not interested resulting in an interest salience $s_{pInt} \in \{0, 1\}$. In case $s_{iLM} = s_{pInt} = 1$ the landmark is interesting for the traveller. Supposing that there is still not one unique landmark, we have to decide if it is beneficial to have more than one suitable landmark available (e.g. for exploratory travel). Otherwise, a decision criterion might be applied (e.g. the object with the shortest distance to the decision point is used).

Travellers with $s_{PspK} \leq 3$ already have familiarity. Thus, semantics are important and a process to determine the object with the maximum semantic salience is included. How detailed this familiarity is depends on the knowledge of the surrounding environment. Therefore, the next decision of the flowchart is if the traveller has landmark, route, or survey knowledge of the surrounding area.

We assume for $s_{PspK} = 3$ that travellers already know some important POIs with semantic and visual salience. Semantic salience is already confirmed for all objects at decision points where the travellers have been before, therefore, a process for the determination of the maximum visual salience follows. In case there is more than one landmark with maximum visual salience, the flowchart proceeds with the interest salience before it reaches a decision.

$s_{PspK} = 2$ means that travellers already know some routes in the environment of the decision point. While passing these routes, their attention might be attracted by structural salient objects. Thus, the flowchart considers structural salience. The interest salience is investigated if there remains more than one landmark.

In case $s_{PspK} = 1$ the traveller is very familiar with the environment. Nevertheless, it is not quite sure whether all available objects are familiar. In addition, we assume that some landmarks are familiar because of their semantic salience but some because of their visual or structural salience. Therefore, the process to identify the object with the highest visual salience is followed by a process to determine the landmark with the maximum structural salience. Supposing there is still more than one landmark left, the decision rule for the interest salience is applied followed by the final decision on the most personalised landmark.

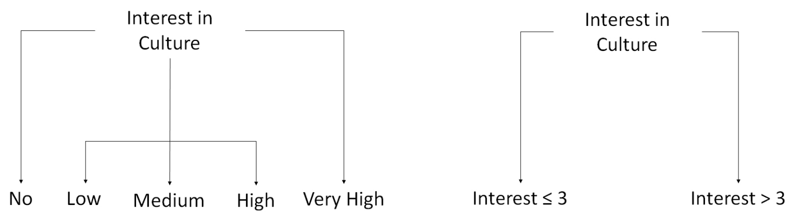
The PdFc provides one or more landmarks for a decision point. Figure 4.6 shows the flowchart. In Section 6.1.2 we investigate and test the PdFc with our training set and - if needed - change the flow.

4.3.7 Conventional Decision Tree Model (CdTm) and Personalised Decision Tree Model (PdTm)

The structures of the CdTm and the PdTm are quite similar. Numerous decision tree algorithms are conceivable as a basis for decision tree models for landmark identification. Which decision tree is the most suitable one depends on the target variable, the values of the attributes, and the general goal. Our general goal is to identify whether an object is a (personalised) landmark or not. Thus, the target variable can take two values either *landmark* (LM) or *not a landmark* (NAL).

The attributes used for the classification are numerical values. Visual, semantic, and structural salience are numeric by default. *Prior spatial knowledge* and *personal interests* ratings, however, could be processed either as numerical or categorical values. Categorical data often require more than two decisions resulting in more than two internal nodes (e.g. we might have one node for every topic of interest and every possible interest rating) which makes it nearly intractable with plenty of possible values. The large number of outcomes is not desirable because the number of data associated with each partition might be too small for any reliable prediction (Tan et al. 2006). One way to overcome such a problem is the restriction to binary splits. Instead of having internal nodes with more than two decisions (Figure 4.7a) we have a binary tree with two decisions (Figure 4.7b). Thus, we treat all attributes as numerical attributes.

There are numerous algorithms for decision tree growing. We prefer an algorithm which is able to handle numerical data and to construct binary trees. We use CART (Breiman et al. 1984), which has been used extensively in the past years (Apté & Weiss 1997). In addition, to meeting all criteria, the algorithm has the advantage that it is not significantly affected by outliers in the input space (Mubayi 2017). This effect is due to the fact that the splitting does not happen on absolute values



(a) Multiway Split. (b) Binary Split.

Figure 4.7: Splitting conditions for different attribute types (modified from Tan et al. (2006)).

but on a proportion of samples within the split ranges (Nisbet et al. 2009). This is particularly useful because there might be objects in our data whose attribute values differ from the attribute values of the other objects (e.g. an object with a higher visual salience than all the others). A further benefit of CART is that it can use the same parameters more than once in different parts of the tree (Nisbet et al. 2009). This capability can reveal complex relationships between sets of parameters. For example, semantic salience might be interdependent on survey knowledge but also on route knowledge. Finally, CART can be used in conjunction with other prediction methods to select the input set of parameters (Nisbet et al. 2009). This is particularly important for decision tree pruning. We use CART together with cross-validation - a pre-pruning method which stops the growing of the tree earlier, before it perfectly fits the training set and, thus, avoids overfitting (Dietterich 1995). The resulting tree is able to classify objects in LM and NAL. We train the CdTm in Section 5.4.4 and the PdTm in Section 6.1.3.

Chapter 5

Data Collection and Preparation

In this chapter we describe the data collection for both, landmark and personal dimensions (Section 5.1). We start with landmark dimensions for objects along an inner city route through Augsburg (Section 5.1.1). The route starts and ends at the Königsplatz, is around 640 meters long, and includes 10 decision points (Figure 5.1). Some of them are famous places of Augsburg such as Moritzplatz and Fuggerplatz. The objects at the decision points consist of 44 buildings, two fountains, and a statue. All personal dimensions are collected by a survey (Section 5.1.2). Section 5.2 discusses the calculation of salience focusing on the collected data for the objects along the route. We describe the division of the collected dataset for the training and testing of the models (Section 5.3). We calculate overall salience measures of the objects and classify objects in *landmark* (LM) and *not a landmark* (NAL) with the help of the conventional models at the end of this chapter (Section 5.4).

5.1 Data Collection for the Dimensions

This section describes the data collection for both landmark and personal dimensions. While landmark dimensions are extracted from official databases or acquired during field surveys, personal dimensions are collected by a survey.

5.1.1 Landmark Dimensions

Identifying landmarks requires attribute data (i.e. visual, semantic, and structural) of the objects as well as information on the corresponding topic of interest. This thesis uses OSM data, official databases, and field survey data. The attributes surface area, height, and colour of the visual dimension are salient only if their values differ from the values of the surrounding objects in a local neighbourhood. Therefore, we collect

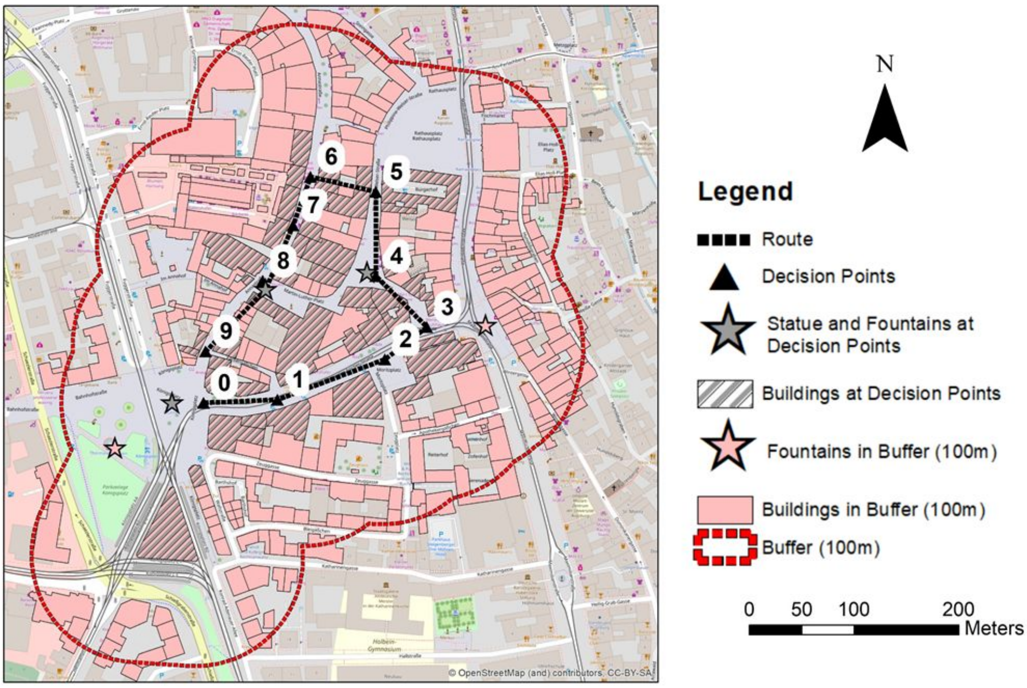


Figure 5.1: Route with decision points.

these attribute values not only for the objects at the decision points, but also for objects in a buffer of a specific size. Following Raubal & Winter (2002) we apply a buffer size of 100 meters (see Figure 5.1).

There are buildings located in backyards within the buffer (Figure 5.2). For these buildings it is not possible to determine colour and surface area. As these objects are not visible from the street, they do not influence salience of the other objects. Therefore, we exclude them from further analysis.

We describe the data sources for the landmark dimensions in the following sections. Figure 5.3 shows a sample landmark.

Visual Dimension

We collect attribute values for the visual dimension during a field survey. In addition, we use OSM data and an official city model as data source.

Surface structure Each single object located at a decision point of the route is investigated on site. Surface structure is one of the attributes directly assessable from the street.

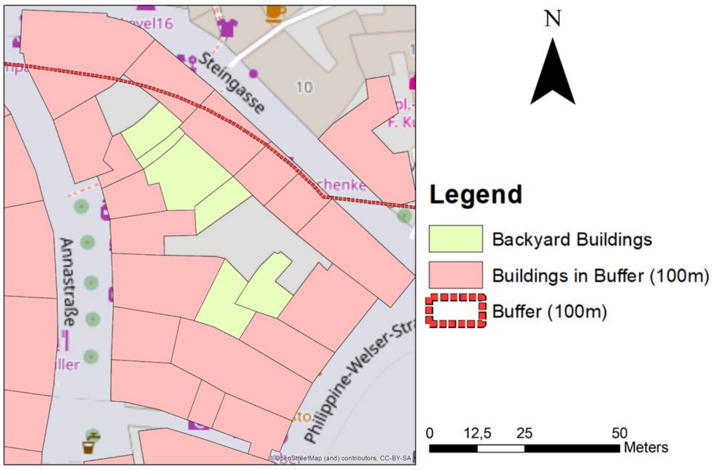
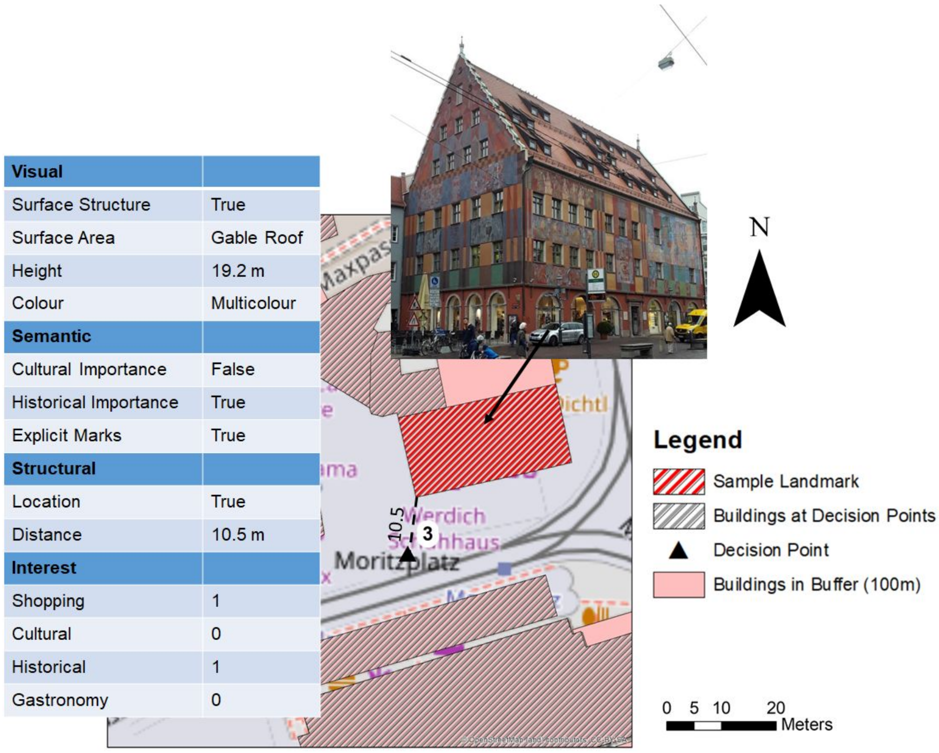
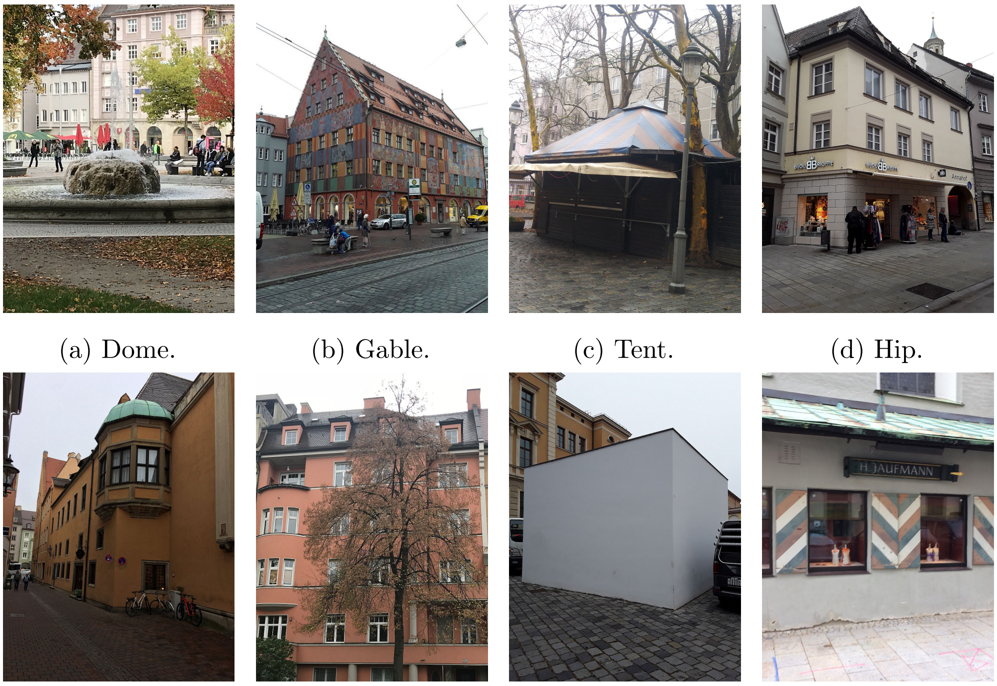


Figure 5.2: Buildings in backyards.



Visual	
Surface Structure	True
Surface Area	Gable Roof
Height	19.2 m
Colour	Multicolour
Semantic	
Cultural Importance	False
Historical Importance	True
Explicit Marks	True
Structural	
Location	True
Distance	10.5 m
Interest	
Shopping	1
Cultural	0
Historical	1
Gastronomy	0

Figure 5.3: Visual (vis), semantic (sem), and structural (str) dimensions as well as landmark interest dimension (shopping (shop), cultural (cult), historical (hist), and gastronomy (gast)) of a sample landmark.



(a) Dome. (b) Gable. (c) Tent. (d) Hip.

(e) Hood. (f) Mansard. (g) Flat. (h) Pent.

Figure 5.4: Different types of surface area.

Surface area We identify the surface area of the objects on-site. We distinguish among several types of surface area (Figure 5.4). We use Google Maps imagery to identify the shape of the roof should it not be visible from the street.

Height An official 3D city model (LOD1 - block model) provides the height of the buildings. We manually estimate the height of the other objects (three fountains and a statue).

Colour There is no official database concerning colour. Hence, we capture the colour of the objects during the field survey.

Semantic Dimension

We collect values for the attributes of the semantic dimension on-site. Additionally, we use an official database.

Cultural importance There are along the route eight objects being classified as cultural important. These are two churches, two fountains, a statue, a museum, and the entrances to the town market.

Historical importance There is an official list of historic monuments in Augsburg (Bayerisches Landesamt für Denkmalpflege 2018). We consider objects that are part of this list as historically important.

Explicit marks Explicitly marked objects must be visible from the street. We capture these on-site.

Structural Dimension

We derive the attributes of the structural dimension from OSM data. OSM provides footprints of the objects (see Figure 5.3). We investigate the topological relations between these footprints and the nearest decision point.

Location at a decision point All objects along the route are located at decision points. Thus, they all receive the Boolean value *True* for that attribute.

Distance to the decision point For each object we calculate the euclidean distance of the nearest point (e.g. the nearest house corner) to the decision point. We use the footprints from OSM (see Figure 5.3).

Landmark Interest Dimension

We analyse the objects in the inner city of Augsburg in order to identify possible topics of interest. This identification is rather subjective and might change dependent on the person who is doing it. We use aids and check the topics of interest with the help of Google Maps and OSM data and validate them with on-site investigations to avoid subjectivity as much as possible. Table 5.1 shows the resulting topics of interest. The eight culturally important objects (see Section 5.1.1, semantic dimension, cultural importance) are also *culturally interesting*. Objects belonging to the topic of interest *historic* are objects from the official list of historic monuments in Augsburg (Bayerisches Landesamt für Denkmalpflege 2018). We classify two buildings hosting a bank as *financially interesting*. One building is a charitable organisation and, therefore, *socially interesting*. We classify the two churches as of *religious interest*. There are 30 *shopping* facilities and 15 places with *gastronomy*, such as bakeries, snack bars and restaurants. We assign the entrances to the town market to *shopping* and *gastronomy* because they are in close connection to all the food offers and the small shops and stands on the market. There is one language school on the route we classify as interesting because of the *educational* character and four buildings with

Table 5.1: Number of objects belonging to different topics of interest at the decision points.

Topic of Interest	Number of Objects	Topic of Interest	Number of Objects
Cultural	**8**	Gastronomy	**15**
Historic	**21**	Sports	0
Supplier	0	Leisure	0
Arts	0	Education	1
Financial	2	Tourism	2
Social	1	Health	4
Religious	2	Nature	0
Shopping	**30**	Architecture	1

medical practices or pharmacies which we classify with *health*. There are a lot of objects which might be interesting for tourists in an inner city area. However, there is no database for touristic monuments and that is why we only classify the building with the tourist information and the station building at the Königsplatz as of *touristic interest* (Figure 5.5). The Königsplatz is an important inner-city transport hub and, therefore, in our eyes interesting for tourists. The topic of interest *architecture* is elusive and difficult to measure because most of the buildings in an inner city's historical area show some outstanding architecturally interesting attributes. We only classify the station building at the Königsplatz as of architectural interest because its appearance is totally different from the other buildings next to the route (Figure 5.5). We do not assign objects to the topic of interest *supplier* since there are no such facilities (e.g. electricity supply companies) along the route. There are no *artificial* objects along the route as well as no *sports* or *leisure* facilities. There are no *natural* objects such as trees or green areas at the decision points.

As Table 5.1 shows most of the objects are of cultural or historical interest, or are shops or gastronomy objects. There are only a few objects available for the other topics of interest. Therefore, we decide to consider only the *personal interests cultural (cult)*, *historical (hist)*, *shopping (shop)*, and *gastronomy (gast)* in our models. Each decision point along the route hosts objects belonging to different topics of interest. Figure 5.3 shows an example object with the topics of interest it belongs to.

Figure 5.5: Architecturally and touristically interesting building at Königsplatz.

5.1.2 Personal Dimensions

We use ESRIs Survey123 for data collection for the personal dimensions. The tool allows to create and publish survey forms (Survey123 2018). This section describes the survey, gives an overview on the participants who completed the survey, and discusses the results.

Participants

One challenging objective was to find a group of participants that is diverse regarding age, education, place of residence, and place of birth. A number of students completed the survey during university lectures. To have participants outside the typical university age and outside the geoinformatics domain, we acquired participants also via personal contacts. In total, 51 people, 24 of whom females, participated in the survey. The average age of the participants is 33.1 ($min = 19$ years, $max = 73$ years, $sd = 15.16$). 23 participants live in Augsburg, 7 of them since their early childhood ($age \leq 10$) or birth. Most of the participants (except six) are born in Germany.

Persons willing to participate had to confirm that they understood that the data collected is used for scientific purposes exclusively. They were told that the data are not forwarded to third parties at any time and that data collection is based on pseudonyms. To this end, they had to confirm that their device's ID will be stored in addition to the data they explicitly enter. In an early version of the survey this confirmation was not included but this survey was completed with students of a university lecture who were notified orally.

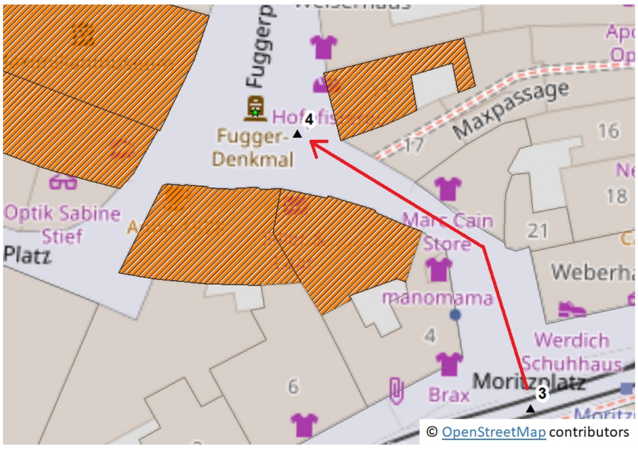


Figure 5.6: Maps in Survey123.

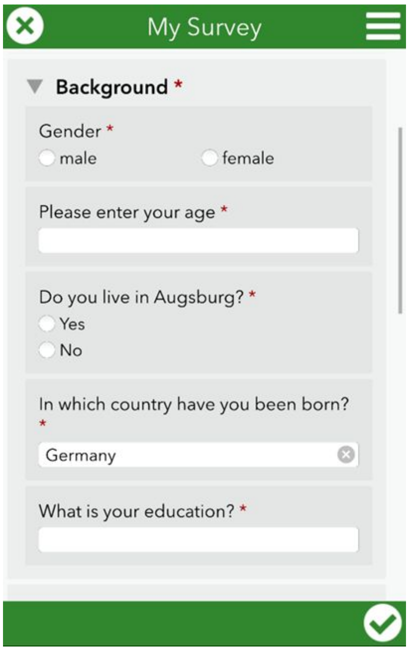
The participants completed the survey between October 2018 and February 2019. We did not collect data in the last week of November and the whole month of December 2018 because of the Christmas market stands. Some of them blocked the view on the objects, especially the stands that were installed around the fountains and the statue.

Procedure

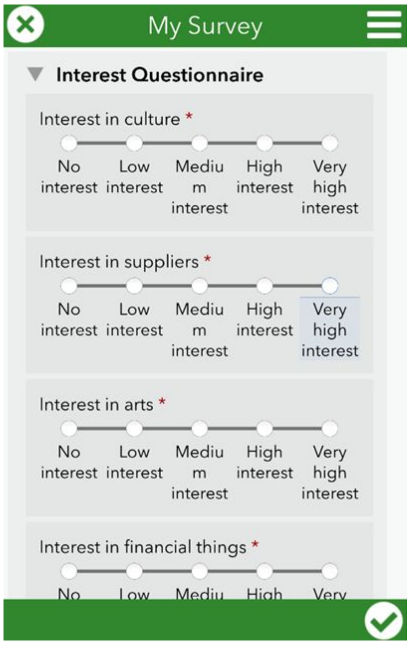
The participants were informed of the starting point of the route. To avoid the influence of turning directions they did not know the whole route in advance. The application guided them from one decision point to the next with the help of maps (Figure 5.6). Most of the participants were guided along the route, some participants completed the survey alone after a comprehensive introduction (installation, procedure, objects to select).

The survey contains a questionnaire focusing on the background of the participants including questions about gender, age, place of residence, and education (Figure 5.7a). It contains questions about *personal interests* (Figure 5.7b), *prior spatial knowledge* (Figure 5.7c), and about objects at the decision points along the route (Figure 5.8).

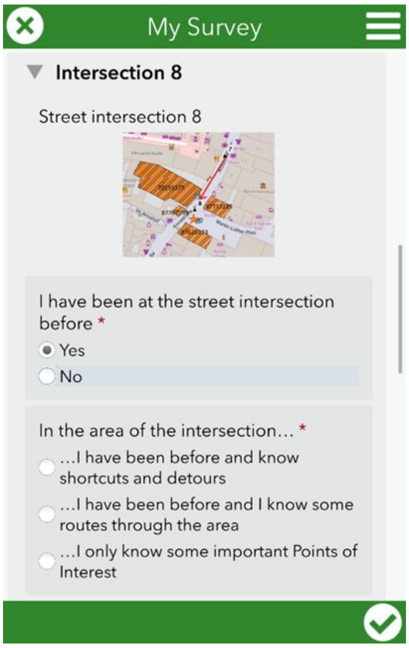
Personal interests The interest questionnaire contains questions about the participant's interest in culture, arts, tourism, historical monuments, nature, architecture, financial things, gastronomy, and facilities of sports, suppliers, leisure, social life, shopping, education, medicine, and religion (Figure 5.7b). Responses to the interest questionnaire are rated on a rating scale with items *no*, *low*, *medium*, *high*, and *very high*.



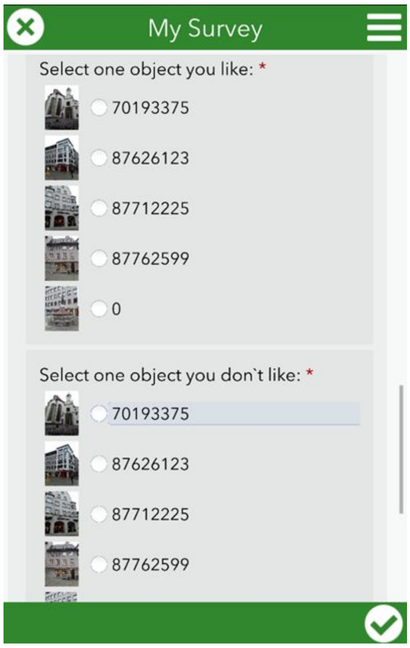
(a) Background questionnaire.



(b) Interest questionnaire.



(c) Spatial knowledge questionnaire.



(d) Object selections.

Figure 5.7: Survey for personal data collection.

Table 5.2: Interest ratings for the personal interests (pInt) on a rating scale (1=no, 2=low, 3=moderate, 4=strong, and 5=very strong interest).

Interest	1	2	3	4	5	Interest	1	2	3	4	5
Cultural	**0**	**6**	**21**	**19**	**5**	**Gastro**	**0**	**2**	**15**	**26**	**8**
Historic	**1**	**6**	**24**	**17**	**3**	Sports	5	10	12	11	13
Supplier	2	16	11	21	1	Leisure	5	2	29	12	3
Arts	0	16	23	7	5	Education	2	13	19	13	4
Financial	6	13	22	10	0	Tourism	2	17	17	12	3
Social	1	6	9	21	14	Health	4	14	18	12	3
Religious	12	19	14	4	2	Nature	0	1	9	20	21
Shopping	**3**	**8**	**12**	**19**	**9**	Architecture	2	6	20	16	7

Prior spatial knowledge At each particular decision point the participants answer questions about their specific spatial knowledge at the individual street intersection and in the area of the street intersection. Based on Table 4.2 participants are first asked if they have been at the street intersection before (Figure 5.7c). According to the response of the participant the survey asks about survey, route, and landmark knowledge (for yes) or survey, route, landmark, or no knowledge (for no) in the area of the intersection (Table 4.2).

Objects at decision points The survey shows a map giving information how to proceed to the next decision point (Figure 5.6). The application additionally shows photos of the objects at the decision points (Figure 5.7d). However, the photos are meant to help the participant to identify the objects in reality. Participants have to look at the real objects to be able to do the selection. We assume that in the case that travellers ask us for route directions, we automatically infer things about the travellers themselves. Therefore, we told participants that they should imagine not common but personally addressed route directions. Based on this assumption they had to select an object they *like* as a *landmark* (LM) and one object they *do not like* (*not a landmark* (NAL)) (Figure 5.7d) for such a route direction. For both questions the same objects are provided. In addition, we ask survey participants to provide a reason for their selections. The survey repeats the procedure for all 10 decision points.

Table 5.3: Numbers of selections for prior spatial knowledge (PspK) (Table 4.2).

s_{PspK}	0	1	2	3	4	5	6	7	8	9	∅
1	21	23	23	24	22	20	21	23	22	20	21.9
2	15	11	12	13	13	14	10	9	11	13	12.1
3	9	9	7	6	8	8	9	9	8	9	8.2
4	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0
6	3	1	1	1	0	0	1	0	1	2	1
7	3	7	8	7	8	9	10	10	9	7	7.8

Results

This section investigates the results of the survey. The outcome is a dataset consisting of *personal interests* ratings, information on *prior spatial knowledge*, as well as objects that participants selected as LM and objects which they selected as NAL for a personally addressed route direction.

Personal interests The survey results in ratings for topics of interest for all 51 participants. As there are only a few objects available for the other topics of interest, we decided to restrict ourselves in this work to the topics of interest shopping, culture, historical monuments, and gastronomy. Nevertheless, for the sake of completeness, we list the other interest ratings (Table 5.2).

Most of the participants showed a *high* interest in gastronomy. Two participants stated that their interest in gastronomy is *low*. There were three participants showing a *very high* interest in historical monuments. Except for seven persons who had *no* or *low* interest in historical monuments, the majority showed a *medium* or even a *high* interest. There were five persons with a *very high* interest in culture. Most of the participants showed a *medium* or a *high* interest in culture. Six stated only a *low* interest. None of the participants stated *no* interest in culture. There were nine people with a *very high* interest in shopping and three with *no* interest. The others rated their interest in shopping somewhere inbetween.

Prior spatial knowledge Participants provided information on their *prior spatial knowledge* at the street intersections. Table 5.3 shows the collected data. On average, 21.9 participants said that they knew the street intersection and that they are familiar with the area. Route knowledge ranked second place (on average 12.1 participants).

Table 5.4: Number of selected objects as landmark (LM)/not a landmark (NAL) at the decision points.

DP	O_1	O_2	O_3	O_4	O_5	O_6
0	16/6	18/3	3/35	10/2	3/4	
1	8/21	14/4	0/2	24/7	4/16	
2	36/0	9/7	4/2	2/23	0/19	
3	5/4	0/17	41/1	1/9	3/19	
4	6/3	6/1	0/9	0/19	0/18	39/1
5	26/6	11/7	12/10	2/28		
6	19/7	8/20	7/12	16/11		
7	3/10	13/11	5/23	27/4		
8	9/7	23/3	1/11	3/25	15/5	
9	12/14	8/4	23/13	8/20		

On average 8.2 participants said that they know the street intersection as well as some important points in the area. Options 4 and 5, implying that a participant who is not at all familiar with the street intersection but has survey or route knowledge of the area, were not chosen at all. On average 1 participant stated that s/he is at this intersection the first time, however knows some important points in the area. The last option, no knowledge of the intersection nor the area was chosen by 7.8 participants on average.

Objects at decision points In total, 47 objects were presented by the survey with a mean of 4.7 (min = 4, max = 6) objects per decision point. We expect that all participants select one object for a LM and one for NAL and that both objects differ. Unfortunately, this was not always the case. As for LMs and NALs the same objects were provided, there were decision points where participants selected the same object for LM and NAL. These decision points are excluded for these participants from further analysis. In total, we collected 503 LMs and the same number of NALs. Table 5.4 lists the number of selections for the LMs and NALs at the decision points. O_3 at decision point 3 is the object most frequently chosen for a LM (Figure 5.8a). In addition, it is selected only once as a NAL. This is not surprising as this object has high visual, semantic, and structural salience. The participants state that they like it primarily because of its colour. People familiar with the intersection and its area state that it is even a famous landmark in Augsburg.

(a) Most frequently as a LM (O_3 at DP 3).

(b) Most frequently as a NAL (O_3 at DP 0).

(c) Selected only as LM (O_1 at DP 2).

(d) Selected only as NAL (O_3 at DP 1).

(e) Selected only as NAL (O_5 at DP 2).

(f) Selected only as NAL (O_2 at DP 3)

(g) Selected only as NAL (O_3, O_4, and O_5 at DP 4).

Figure 5.8: Object selections as LMs and NALs.

Completely different: O_3 at decision point 0 is most frequently selected as a NAL (Figure 5.8b). This object changes its appearance during summer and winter - in summer it is a fountain with a figure in winter it is within a box. It has no semantic salience and is only visually salient in height. Participants state that the box is too inconspicuous, small and not well visible, and some of them did not even know what it is. There is one object selected only as LM (Figure 5.8c, decision point 2, O_1). The participants did not select the objects in Figures 5.8d - Figure 5.8g as LMs.

5.2 Calculating Salience for the Dimensions

In this section, we discuss briefly how salience values are calculated for landmark as well as for personal dimensions from the collected data. In addition, we present the resulting datasets which are used as input datasets for the personalised landmark identification models.

5.2.1 Landmark Dimensions

We calculate the salience values for the attributes of landmark dimensions according to Table 4.1. Then we group the salience values for the visual, the semantic, and the structural dimension. This results in percentage values for s_{vis}, s_{sem}, and s_{str}.

We check each object whether it belongs to the topics of interest shopping, culture, historical monuments, or gastronomy. In case the object belongs to a specific topic of interest we classify it as salient for this topic. Thus, we need no further processing for obtaining s_{iLM}.

5.2.2 Personal Dimensions

The survey is interest-oriented and allows to rate it on a five point rating scale. The PdFc distinguishes between interested and not interested (Section 4.3.6). We consider the ratings *no*, *low*, and *medium* as *not interested* and *high* and *very high* ratings as *interested*. This results in a salience of $s_{pInt} \in \{0, 1\}$ for the PdFc. For the other models we consider the original interest ratings ($s_{pInt} \in \{1, ...5\}$).

Survey participants rate their *prior spatial knowledge* on a rating scale with values between 1 and 7 (Table 5.3). We need no further calculations to obtain the salience s_{PspK} and use the ratings of the participants directly in our models.

5.2.3 Input Data for the Models

The results of the salience calculations are the input data for the landmark identification models. The P/CwSm, the P/CwPm, and the P/CdFc only use the LMs for the training (Table 5.5 and Table 5.6). The P/CdTm uses both, LMs and NALs and, therefore, needs this information in the input data (Table 5.7).

Table 5.5: Input data for the P/CwSm and P/CwPm.

DP	ID	Landmark			Personal				
		s_{vis}	s_{sem}	s_{str}	s_{PspK}	s_{pInt}			
						shop	cult	hist	gast
2	O_1	25	50	50	5	4	2	3	5
3	O_3	75	75	100	5	4	2	3	5

Table 5.6: Input data for the P/CdFc.

DP	ID	Landmark							Personal				
		s_{vis}	s_{sem}	s_{str}	s_{iLM}				s_{PspK}	s_{pInt}			
					shop	cult	hist	gast		shop	cult	hist	gast
2	O_1	25	50	50	0	1	1	0	5	1	0	0	1
3	O_3	75	75	100	1	0	1	0	5	1	0	0	1

Table 5.7: Input data for the P/CdTm.

DP	ID	Landmark			Personal					LM/NAL
		s_{vis}	s_{sem}	s_{str}	s_{PspK}	s_{pInt}				
						shop	cult	hist	gast	
2	O_1	25	50	50	5	4	2	3	5	LM
2	O_5	25	0	50	5	4	2	3	5	NAL
3	O_2	0	50	50	5	4	2	3	5	NAL
3	O_3	75	75	100	5	4	2	3	5	LM

The input data for the conventional models contain only information on landmark dimensions, whereas the data for the personalised models additionally include the personal dimensions. Personal dimensions are *prior spatial knowledge* at the particular street intersection and *personal interests* ratings for shop, cult, hist, and gast. The PwSm, the PwPm, and the PdTm consider the original interest ratings of the participants of the survey with values from one to five. As the PdFc distinguishes only between interested and not interested, the *personal interests* of the participants are expressed with zeros and ones in its input data (Table 5.6). In addition, PdFc needs information on the assignment of the objects to topics of interest, which is an additional dimension of the landmark dimensions.

5.3 Data Division in Training and Test Set

We divide our collected dataset into a training set and a test set. We use the training set to train the machine learning models both CdTm and PdTm. The conventional models based on theory (CwSm, CwPm, and CdFc) have no unknown model parameters, whereas the weights of the PwSm and the PwPm and the flow of the PdFc as well need to be identified with the help of the training set. After the training we investigate their performance with the test set. There are three freely adjustable model parameters for the PwSm and the PwPm (Section 6.1.1) and five model parameters for the PdTm (Section 6.1.3). The PdFc does not have model parameters as it is built on decisions and processes. Hence, the training/testing ratio should be in the range of $1/\sqrt{3}$ and $1/\sqrt{5}$ (Section 3.2.3). For reasons of comparability we use a 50:50 training/testing ratio for all models.

There are several options to divide the dataset in training and test set (Section 3.2). Independently collected data is not an option in our case, because we use only one test route. We might split our dataset according to the months in which the survey was completed. However, temporal autocorrelation might lead to dependent training and test sets. This, in turn, might lead to overly optimistic identification of landmarks (Bahn & McGill 2013). We intend to apply the models to identify landmarks also in new geographic spaces, thus, we are required to use spatially independent training and test sets (Bahn & McGill 2013). We choose two sets that do not overlap spatially. We use 50% of the data to train the models and the remaining 50% to test their performance. We divide our dataset consisting of data for the 10 decision points into two sets of equal size: the first five decision points (0 - 4) belong to the training set and the other five (5 - 9) to the test set (Figure 5.9).

The training and test sets for the P/CdTm differ from the ones for the P/CwSm,

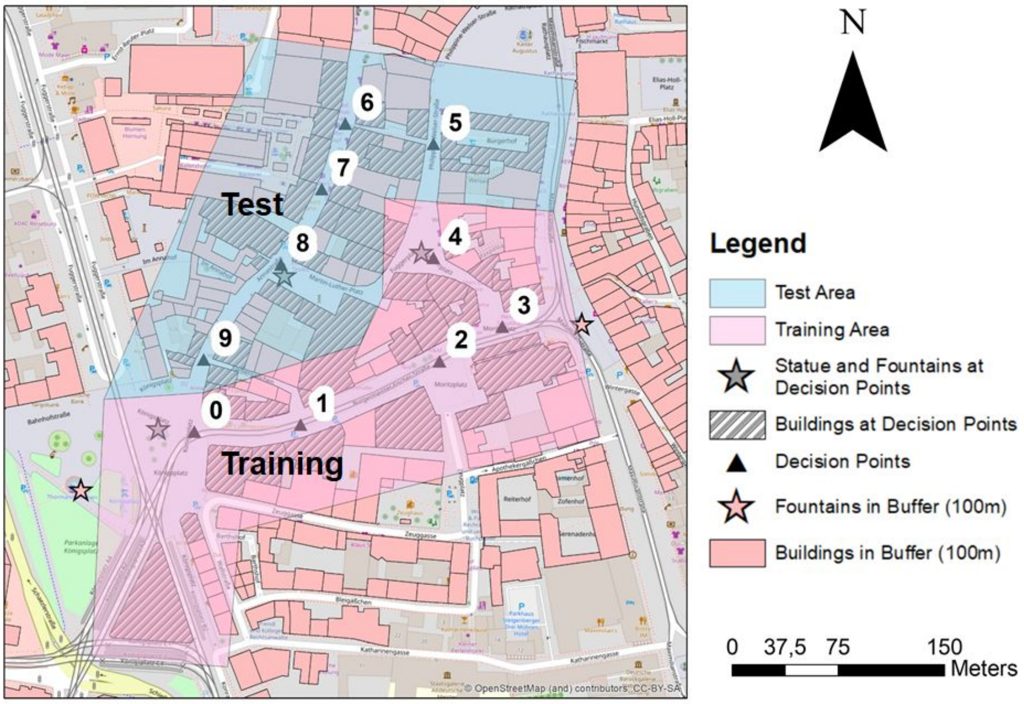


Figure 5.9: Training and test area.

P/CwPm, and P/CdFc. The training set for the P/CdTm includes 252 LMs and 252 NALs, while the training sets for the other models include only 252 landmarks. For PwSm and PwPm it is important that the *prior spatial knowledge* and *personal interests* ratings appearing in the test set also appear in the training set (Section 6.1.1). Therefore, we exclude landmarks with a combination of *prior spatial knowledge* and *personal interests* ratings not appearing in the training set from the test set. The resulting test set consists of 232 landmarks. We do not need NALs for testing (Section 5.2.3) since we are only interest in the identification of landmarks.

5.4 Overall Salience according to Conventional Models

This section discusses the identification of landmarks with the conventional models. We use the results of the conventional models for the comparison and the assessment of the results of the personalised landmark identification models (Section 7.2). The CwSm, the CwPm, and the CdFc have no unknown model parameters, whereas the model parameters of the CdTm need to be identified (trained). For the training of the CdTm and testing of all the conventional models we use the datasets presented in Section 5.2.3.

Table 5.8: Recalls of the conventional models obtained with the training set (accuracy in brackets for the CdTm).

	CwSm	CwPm	CdFc	CdTm
Recall [%]	66.27	60.71	60.71	56.75 (73.61)

Table 5.9: Recalls of the conventional models obtained with the test set.

	CwSm	CwPm	CdFc	CdTm
Recall [%]	40.95	40.95	31.46	67.67

5.4.1 Conventional Weighted Sum Model

We identify landmarks with the CwSm proposed by Raubal & Winter (2002) (Section 4.3.1). We need no training to identify optimal weights because the CwSm is based on theory and weights are set to one ($w_{vis} = w_{sem} = w_{str} = 1$). Thus, we apply the CwSm directly to the training and the test set. In case we apply the CwSm to the whole dataset the recall is 54.13%. The recall on the training set is 66.27% (Table 5.8). For the test set the recall is lower and reaches only 40.95% (Table 5.9). Out of the 232 test set landmarks the CwSm correctly identifies 95. Figure 5.10 (upper left) shows the identified landmarks. The model identifies at least one landmark at each decision point ($n = 17$). The average number of landmarks at a decision point is 1.7 ($min = 1$ (intersections 2, 3, 7, and 9), $max = 3$ (intersection 4)).

5.4.2 Conventional Weighted Product Model

The CwPm considers weights of one ($w_{vis} = w_{sem} = w_{str} = 1$) and needs no training because it is based on theory (Section 4.3.3). The recall on the whole dataset is 51.23%. The recall on the training set is with 60.71% again higher as the one obtained with the test set (40.95%). Figure 5.10 (upper right) shows the landmarks identified with the CwPm applied to the test set. The model identifies at least one landmark at each decision point ($n = 15$). It determines either one or two landmarks for a decision point ($mean = 1.5$, $min = 1$ (intersection 1, 2, 3, 7, and 9) and $max = 2$ (intersection 0, 4, 5, 6, and 8)). The CwPm identifies 95 of 232 landmarks correctly for the test set.

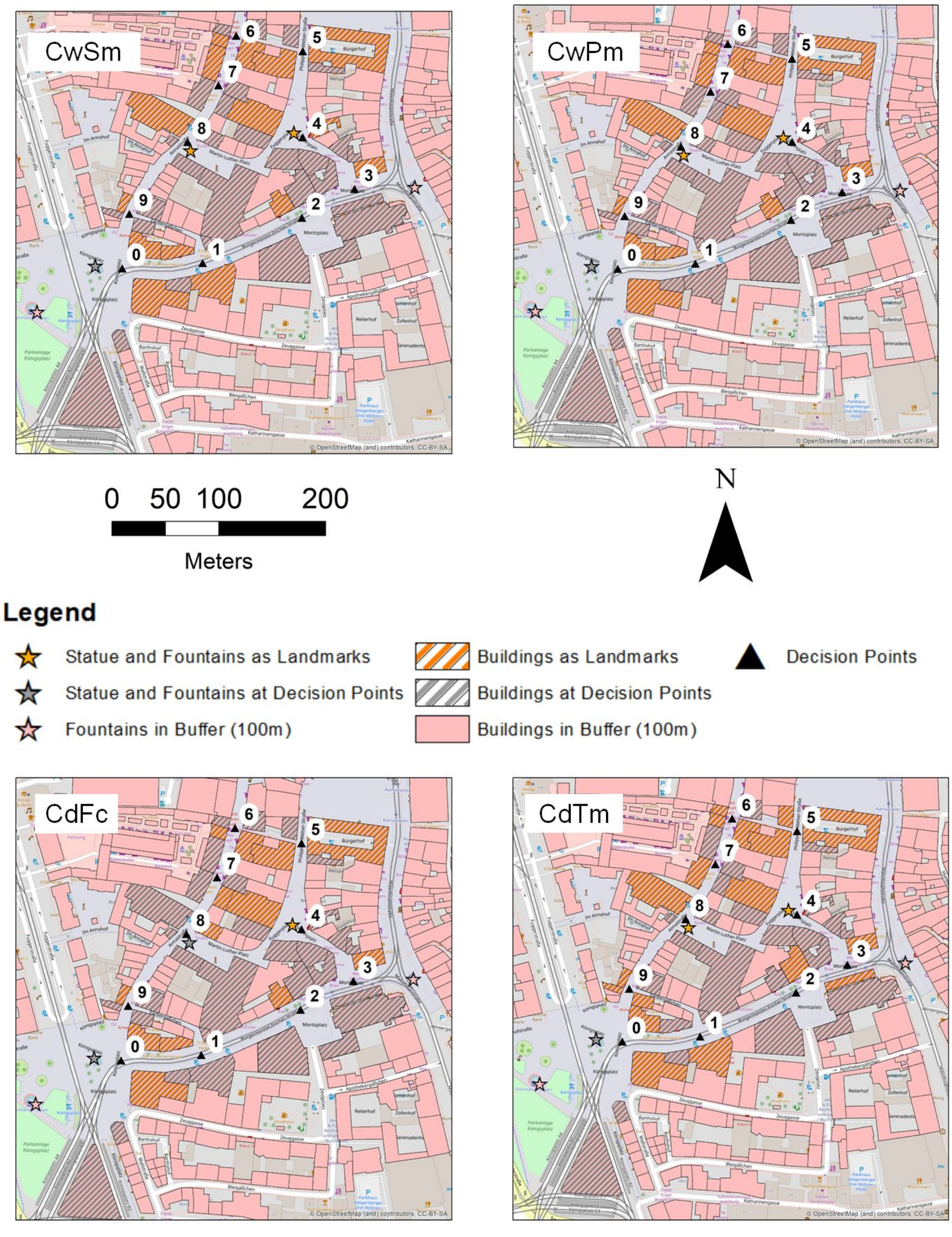


Figure 5.10: Landmarks according to conventional models.

5.4.3 Conventional Decision Flow Chart

For the CdFc we follow the flow chart developed in Section 4.3.5 (Figure 4.5). The recall on the whole dataset is 46.69%. We reach a recall of 60.71% for the training set, whereas the recall of the test set is only 31.46%. This corresponds to 73 correctly identified landmarks. Figure 5.10 presents the results of the CdFc in the lower left. The average of the landmarks is 1.4 ($n = 14$, $min = 1$ (intersection 1, 2, 3, 7, 8, and 9) and $max = 2$ (0, 4, 5, and 6)).

5.4.4 Conventional Decision Tree Model

In Section 4.3.7 we identify CART as suitable for our purposes. The CdTm learns its behaviour from the training set. Section 6.1.3 describes the exact training method for the PdTm. We apply the same method for the CdTm. The Appendix contains the training parameters (Table A.1 and Table A.2). We present the results of the CdTm at this point (Figure 5.10, lower right) without going further in details. The CdTm identifies a total of 20 landmarks ($min = 1$ (intersection 1, 2, 7), $max = 3$ (intersection 5, 6, 9)). It identifies 300 landmarks for the whole dataset (61.98%). Out of the 232 test set landmarks the CdTm identifies 157 correctly which equals 67.67%. The recall for the training set is lower (56.75%) than the accuracy (73.61%) (compare Formula 3.5 and Formula 3.6).

5.4.5 Results of the Conventional Models Discussed

The recalls that the CwSm, the CwPm, and the CdFc obtain on the whole dataset is around 50%. The results for the training set are around 60% (Table 5.8), whereas the recalls for the test set merely reach around 40% (Table 5.9). The CdFc reaches a recall of only 31.46% for the test set. The question arises whether training and test set are well-chosen. In case we reverse the training and test set, decision points 0 - 4 become the test set and decision points 5 - 9 the training set. In this case the CwSm, the CwPm, and the CdFc deliver a recall of around 40% for the training set and a better recall of 60% for the test set. We might modify the training and test set in such a way that they achieve a recall of 50% for both training and test set. This is because the models are theoretically constructed and do not learn their behaviour from the training set. Therefore, we cannot expect any other results from our dataset, except that the obtained recalls for the training and the test set might shift. We do not modify the training and test set and proceed with the specified ones.

The weights of the CwSm and the CwPm are set to one. There are studies saying that different landmark dimensions have a different impact on successful landmark identification which outlines the importance of weighting each dimension relative to its significance (Kattenbeck 2016, Sadeghian & Kantardzic 2008). However, the question whether the recall of the CwSm and CwPm might be improved by considering weights remains open at this point but is discussed in Section 8.3.1.

We build a basic CdFc delivering a similar or even identical training recall as the CwSm and the CwPm respectively. However, the recall on the test set is with 31.46% the lowest and not half as high as the recall of the CdTm. We do not learn the flow from the training set. However, when varying the flow from Figure 4.5 by changing the process of visual and semantic salience, we do not get better results.

The CdTm uses information from LMs as well as from NALs for training. It obtains an accuracy of 73.61% with the training set (Table 5.8, in brackets). The result differs, however, for the recall (56.75%, Table 5.8). The reason for this might be found in the tree (Figure A.11). The CdTm shows a terminal node of the class *NAL* having 97 samples of the class *LM* and 148 samples of the class *NAL*. This is not a *pure* terminal node at all because it shows a gini-index of 0.478 (Section 3.1.2). As this terminal node is declared as the class *NAL*, a number of objects which are actually selected as landmarks end up in this node and are consequently identified as NALs. However, the training of the CdTm with cross-validation and grid-search identifies the model parameters of the CdTm in Table A.2 as the ones yielding the highest average accuracy. Therefore, we continue with the CdTm built on these model parameters. The recall of the CdTm obtained with the test set is higher (Table 5.9) than the one obtained with the training set. This means the CdTm is better able to identify the landmarks in the test set than in the training set.

Table 5.10 shows the results of a McNemar's test applied to the model results on the test set (Section 3.4). The difference of the CdTm to the other three models is considered to be statistically significant and is $p \leq 0.0001$ for all cases. We are not able to calculate a McNemar's test statistic and a p-value for the comparison of the CwSm and the CwPm because no landmarks changed from unidentified to correctly identified or vice-versa. The comparisons between the CwSm and the CdFc or the CwPm and the CdFc respectively show that the difference is extremely statistically significant with a $p < 0.0001$. There are 30 discordant pairs when comparing the CwSm or the CwPm with the CdTm. There are 26 pairs where the CdTm correctly identifies a landmark but CwSm/CwPm does not, and 4 pairs where CwSm/CwPm correctly identifies a landmark but CdTm does not.

Table 5.10: Results of McNemar's Test for the conventional models.

	CwPm	CdFc	CdTm
CwSm			
Identified → Unidentified	0	65	4
Unidentified → Identified	0	3	26
Test Statistic	NaN	54.72	14.7
p-Value	0	<0.0001	0.0001
CwPm			
Identified → Unidentified		65	4
Unidentified → Identified		3	26
Test Statistic		54.723	14.7
p-Value		<0.0001	0.0001
CdFc			
Identified → Unidentified			4
Unidentified → Identified			88
Test Statistic			74.88
p-Value			<0.0001

Summarising the above, we conclude that the CdTm delivers the highest recall and its results differ significantly from the results of the other models. One reason for this might be that the model identifies more landmarks than the other conventional models ($n = 20$). As a consequence it identifies more landmarks selected by survey participants. Another reason for this behaviour might be that the model is not based on theoretical considerations but learns from the training set. For a further discussion on this topic see Section 6.3.

Chapter 6

Training and Testing of the Personalised Models

This chapter describes the creation of the personalised landmark identification models and the subsequent identification of personalised landmarks. In Section 5.3 we divided the collected dataset in training and test set. In Section 6.1 we train the machine learning model on the training set using the traditional approach presented in Section 3.2.1 and we 'train' the models based on theory with an approach inspired by this traditional approach (Section 3.2.2). Subsequently, we use the created models to identify the landmarks of the test set and compare the identified landmarks with the landmarks selected by the participants of the survey (Section 6.2). We close this chapter with a discussion of the results of the training and testing (Section 6.3).

6.1 Training of the Models

In this section we train the personalised landmark identification models with the collected data. The models based on theory, the PwSm, the PwPm, and the PdFc, only use *landmarks* (LMs) for training whereas the machine learning model (PdTm) needs also information on objects, which are *not a landmark* (NALs). The PwSm and the PwPm calculate a salience measure, whereas the PdTm and the PdFc classify objects as LMs and, in the case of the PdTm, NALs.

We use the methods proposed for *'training' of models based on theory* (Section 3.2.2) to identify model parameters for the PwSm and the PwPm and an optimal flow for the PdFc. For the PdTm we use the *traditional machine learning approach* to identify the model parameters (Section 3.2.1). In the following sections we evaluate the results of the model training.

6.1.1 Personalised Weighted Sum Model and Personalised Weighted Product Model

We use the methods proposed for *'training' of models based on theory* (Section 3.2.2, Figure 3.4) to identify model parameters for the PwSm and the PwPm. The first step for the training of the PwSm and the PwPm is to specify initial weights. This section first describes the search for such weights and subsequently, investigates individual results of the PwSm and the PwPm.

Finding Initial Weights

The PwSm and the PwPm have three freely adjustable model parameters: w_{vis}, w_{sem}, and w_{str} (Formulas 4.2 and 4.4). They reflect the influence of the traveller's *personal interests* (pInt) and the *prior spatial knowledge* (PspK) on the personal salience of an object. The first step is to find initial weights for w_{vis}, w_{sem}, and w_{str}. We analyse the objects selected as landmarks by survey participants with different pInt and PspK ratings. For each combination we determine the average of visual, semantic, and structural salience ($\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$, Table A.3). Column *No* in the table shows that some combinations of PspK and pInt ratings appear only once in the training set. In Section 7.2.5 we investigate how this affects the identification of personalised landmarks.

We use initial relative weights to train the PwSm and the PwPm. The minimum value of $\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$ (Table A.3) is used as a reference value to calculate the relative weights:

$$\begin{aligned} min_{\overline{s}} = min(\overline{s}_{vis}, \overline{s}_{sem}, \overline{s}_{str}) \\ w_{visRel} = \frac{\overline{s}_{vis}}{min_{\overline{s}}} \\ w_{semRel} = \frac{\overline{s}_{sem}}{min_{\overline{s}}} \\ w_{strRel} = \frac{\overline{s}_{str}}{min_{\overline{s}}} \end{aligned} \tag{6.1}$$

It might be that either the PwSm or the PwPm do not fit the data with these initial relative weights and, therefore, obtain a low recall (Formula 3.6). For this reason we introduce model parameters p_{vis}, p_{sem}, and p_{str}, and multiply them with the initial relative weights (Formula 6.2).

$$\vec{w} = \begin{pmatrix} p_{vis} * w_{visRel} \\ p_{sem} * w_{semRel} \\ p_{str} * w_{strRel} \end{pmatrix} \tag{6.2}$$

We set the initial relative weights manually in advance before the training starts dependent on the PspK and the pInt ratings. Then we build different personalised PwSms and PwPms respectively with different model parameters. Following the method proposed in Section 3.2.2 we calculate the average recalls of the validations folds for each combination of model parameters. The following sections describe the details for the PwSm and the PwPm.

Training Personalised Weighted Sum Model

For the determination of optimal model parameters for the PwSm we start with a coarse grid-search setting $p_{vis} = p_{sem} = p_{str} = 1$ and increase them alternately by 0.5 until 10. We calculate the average recall of the 10 validation datasets checking each combination of model parameters. We obtain average recalls varying between 39.03% and 62.30%. There are several combinations obtaining the highest average recall thereof, $p_{vis} = 2$, $p_{sem} = 1$, and $p_{str} = 1$ is the one with the smallest values. Table 6.1 shows that the neighbouring combinations of model parameters deliver a lower recall.

Table 6.1: Average recalls for initial coarse grid-search PwSm.

p_{vis}	p_{sem}	p_{str}	Average Recall [%]
1.5	10	10	40.53
2	**1**	**1**	**62.30**
2	1	1.5	43.33

As these recalls are much lower we do not expect to obtain a better average recall with a finer grid-search. However, for the finer grid-search we vary the model parameters around the best values of the coarse search. We start with $p_{vis} = 1.9$, $p_{sem} = 0.9$, and $p_{str} = 0.9$ and increase the model parameters alternately by 0.1 until $p_{vis} = 2.1$, $p_{sem} = 1.1$, and $p_{str} = 1.1$. Table 6.2 shows an extract of the results. The finer grid-search confirms the result of the coarse search.

We build the final PwSm with the model parameters obtained:

Table 6.2: Average recalls for finer grid-search PwSm.

p_{vis}	p_{sem}	p_{str}	Average Recall [%]
1.9	0.9	0.9	60.30
...	...	...	...
2.0	0.9	1.1	55.56
2.0	1.0	0.9	61.07
2.0	1.0	1.0	62.30
2.0	1.0	1.1	58.36
2.0	1.1	0.9	61.10
2.0	1.1	1.0	60.30
...	...	...	...
2.1	1.1	1.1	59.56

$$\begin{aligned} s_{PwSm} &= (p_{vis} * w_{visRel} * s_{vis} + p_{sem} * w_{semRel} * s_{sem} + p_{str} * w_{strRel} * s_{str})/100 \\ &= (2 * w_{visRel} * s_{vis} + 1 * w_{semRel} * s_{sem} + 1 * w_{strRel} * s_{str})/100 \qquad (6.3) \end{aligned}$$

The recall on the given training set is 62.30% (Table 6.5). In Section 6.2.1 we apply the PwSm to the test set and elaborate on the results.

Training Personalised Weighted Product Model

The PwPm has the same three freely adjustable parameters as the PwSm. Similarly to the approach used for the PwSm we start a coarse grid-search with $p_{vis} = p_{sem} = p_{str} = 1$ and increase the values alternately by 0.5 until 10. The obtained average recalls vary between 38.93% and 60.70%. The best recall appears with more than one combination of p_{vis}, p_{sem}, and p_{str}. Table 6.3 shows the combination of model parameters with the minimum values obtaining the best average recall.

Table 6.3: Average recalls for initial coarse grid-search PwPm.

p_{vis}	p_{sem}	p_{str}	Average Recall [%]
1	1.5	10	38.93
1	**2**	**1**	**60.70**
1	2	1.5	53.13

Table 6.4: Average recalls for finer grid-search PwPm.

p_{vis}	p_{sem}	p_{str}	Average Recall [%]
0.9	1.9	0.9	60.70
0.9	1.9	1.0	60.70
0.9	1.9	1.1	59.53
0.9	2.0	0.9	60.70
0.9	2.0	1.0	60.70
0.9	2.0	1.1	59.90
0.9	2.1	0.9	60.70
...	...	...	...
1.0	2.0	1.0	60.70
...	...	...	...
1.1	2.1	1.1	60.70

Table 6.3 shows that the neighbouring values from initial coarse grid-search deliver lower average recalls. Therefore, we do not expect better values from a finer grid-search. However, similarly as for the PwSm, we vary the model parameters around the best values of the coarse search. We start with $p_{vis} = 0.9$, $p_{sem} = 1.9$, and $p_{str} = 0.9$ and increase the values alternately by 0.1 until $p_{vis} = 1.1$, $p_{sem} = 2.1$, and $p_{str} = 1.1$. Table 6.4 shows an extract of the result.

For most of the parameter combinations the model results in a 60.70% recall on average. With two exceptions for $p_{vis} = 0.9$, $p_{sem} = 1.9$, and $p_{str} = 1.1$ and $p_{vis} = 0.9$, $p_{sem} = 2.0$, and $p_{str} = 1.1$ resulting in slightly lower recalls.

We use the minimum model parameters expressed with a whole number as model parameters for the PwPm. We build the model as follows:

$$
\begin{aligned}
s_{PwPm} &= (s_{vis}^{p_{vis}*w_{visRel}} * s_{sem}^{p_{sem}*w_{semRel}} * s_{str}^{p_{str}*w_{strRel}})/100 \\
&= (s_{vis}^{1*w_{visRel}} * s_{sem}^{2*w_{semRel}} * s_{str}^{1*w_{strRel}})/100 \qquad (6.4)
\end{aligned}
$$

The recall on the given training set is 60.71% (Table 6.5). We evaluate the performance of our model on the test set in Section 6.2.2 .

Table 6.5: Recalls of the personalised models obtained with the training set (accuracy in brackets for the PdTm).

	PwSm	PwPm	PdFc	PdTm
Recall [%]	62.30	60.71	64.68	78.97 (78.17)

6.1.2 Personalised Decision Flow Chart

The PdFc does not have model parameters because it is built on decisions and processes. Hence, we vary the flow of the model and use the methods proposed for *'training' of models based on theory* (Section 3.2.2, Figure 3.4) to identify an optimal flow for the PdFc. The average recall for the flow in Figure 4.6 is 60.29% (Table 6.6). We investigate whether it is possible to achieve a higher recall and train the model with the following modifications:

- Adjust the flow for $s_{PspK}(Intersection) > 3$ and skip $max(s_{str})$.
- Adjust the flow for $s_{PspK}(Intersection) > 3$ and skip $max(s_{vis})$.
- Adjust the flow for $s_{PspK}(Intersection) \leq 3$ and follow $s_{PspK} = 3$.
- Adjust the flow for $s_{PspK}(Intersection) \leq 3$ and follow $s_{PspK} = 2$.
- Adjust the flow for $s_{PspK}(Intersection) \leq 3$ and follow $s_{PspK} = 1$.
- Adjust the flow for $s_{PspK}(Intersection) \leq 3$ and skip $max(s_{vis})$ and $max(s_{str})$.

The modifications result in several combinations. Table 6.6 shows their average recalls for the 10 validation folds. The best average recall is obtained by the adjusted flow for $s_{PspK}(Intersection) > 3$ *and skip* $max(s_{str})$ together with the adjusted flow for $s_{PspK}(Intersection) \leq 3$ *and skip* $max(s_{vis})$ *and* $max(s_{str})$. This flow obtains an average recall of 64.69%. Figure 6.1 shows the adapted flow chart. When calculating the results for this flow, we obtain a recall of 64.68% in the training set (Table 6.5).

We tested a flowchart which considers first the decision $s_{iLM} = s_{pInt} = 1$ and only afterwards the decisions on s_{PspK} (Figure A.12). Table A.4 shows an extract of the average recalls. However, we reject this option because the average recalls are lower than the other way round. In Section 6.2.3 we test how many landmarks the flow in Figure 6.1 is able to identify for the test set.

Table 6.6: Average recalls for different flows.

$s_{PspK}(Intersection)$		Average Recall [%]
>3	≤ 3	
see Figure 4.6	see Figure 4.6	60.29
skip $max(s_{str})$	see Figure 4.6	63.89
skip $max(svis)$	see Figure 4.6	56.36
see Figure 4.6	flow $s_{PspK} = 3$	60.29
skip $max(s_{str})$	flow $s_{PspK} = 3$	63.89
skip $max(s_{vis})$	flow $s_{PspK} = 3$	56.36
see Figure 4.6	flow $s_{PspK} = 2$	60.29
skip $max(s_{str})$	flow $s_{PspK} = 2$	63.89
skip $max(s_{vis})$	flow $s_{PspK} = 2$	56.36
see Figure 4.6	flow $s_{PspK} = 1$	60.29
skip $max(s_{str})$	flow $s_{PspK} = 1$	63.89
skip $max(s_{vis})$	flow $s_{PspK} = 1$	56.36
see Figure 4.6	skip $max(s_{vis})$ and $max(s_{str})$	61.10
skip $\boldsymbol{max(s_{str})}$	**skip $\boldsymbol{max(s_{vis})}$ and $\boldsymbol{max(s_{str})}$**	**64.69**
skip $max(s_{vis})$	skip $max(s_{vis})$ and $max(s_{str})$	57.16

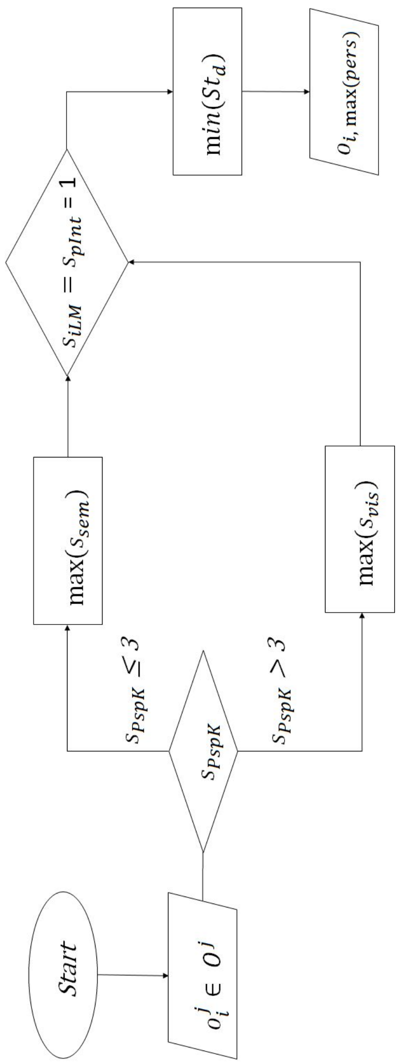


Figure 6.1: Final Personalised Decision Flow Chart.

6.1.3 Personalised Decision Tree Model

We use the *traditional machine learning approach* (Section 3.2, Figure 3.3) to identify the model parameters of the PdTm. It has several freely adjustable model parameters that can be tuned to optimise the identification result. The most common ones are (Scikit 2018, Pedregosa et al. 2011):

- **Criterion** The function which measures the quality of the split. It can be *gini* or *entropy* (Section 3.1.2).
- **Splitter** The method used to select the split at each node. In case *best* is selected, the tree splits on the most relevant feature. In case of *random*, the tree takes a random feature and splits it.
- **min_samples_split** The minimum number of samples required to split a tree node. A split is not performed as soon as there are less than a certain number of samples.
- **min_samples_leaf** The minimum number of samples required to be at a leaf, at the base of the tree.
- **max_depth** The maximum depth of the tree indicates how deep the tree can grow. The depth is the length of the longest path from the root node to a leaf. It captures more information the deeper it is and the more splits it has.

There are model parameters giving the opportunity to weight things higher than others (e.g. the target values LM or NAL or e.g. a specific PspK or a pInt). However, we give none of them a higher weight, therefore, PspK and pInt have equal weights and the target values are supposed to have weight one (Scikit 2018). The number of dimensions to consider when looking for the best split might be considered to train the tree. As we do not want to restrict the possible results, we use all dimensions and perform no attribute subset that could be selected during decision tree growing.

We use grid-search with cross-validation to identify optimal model parameters (Section 3.2). For the PdTm we need training folds as well as the validation folds because the model learns from the training set (Figure 3.3). Table 6.7 shows the initial model parameter settings for the coarse grid-search. We evaluate and compare the results of the cross-validation looking at the model parameters obtaining the highest average accuracy. We identify the highest average accuracy with 76.78% for the model parameters shown in Table 6.7.

Table 6.7: Parameter values for initial coarse grid-search PdTm.

Parameter	Value	Best Value
Criterion	gini, entropy	entropy
Splitter	best, random	random
min_samples_split	[5, 10, ..., 50]	30
min_samples_leaf	[5, 10, ..., 50]	5
max_depth	[5, 10, ..., 50]	10
Average Accuracy [%]	-	76.78

Table 6.8: Parameter values for finer grid-search PdTm.

Parameter	Value	Best Value
Criterion	gini, entropy	gini
Splitter	best, random	random
min_samples_split	[25, 26, ..., 35]	34
min_samples_leaf	[1, 2, ..., 10]	5
max_depth	[5, 6, ..., 15]	9
Average Accuracy [%]	-	77.38

In a next step we conduct a finer grid-search, varying the values of *min_ samples_ leaf*, *min_ samples_ split*, and *max_ depth* around their best values obtained by the coarse grid-search. Table 6.8 shows the best average accuracy with the model parameters of the finer search. After we found the best parameters, we train the PdTm on the training set to generate the final classifier. Figure 6.2 shows the resulting tree.

The nodes and leaves of the PdTm are coloured by their class (orange = LM, blue = NAL). They indicate the splitting criterion used, namely the gini-index (Section 3.1.2). The intensity of the colour gives information on the height of the gini-index. In the root node of the PdTm (Figure 6.2) the probability of obtaining two different outputs is 0.5. The tree shows in the left part a terminal node with $gini = 0$. This is a *pure* terminal node because at this point the tree always identifies the object as a NAL. This means a 100% accuracy in identifying the right class for the training data.

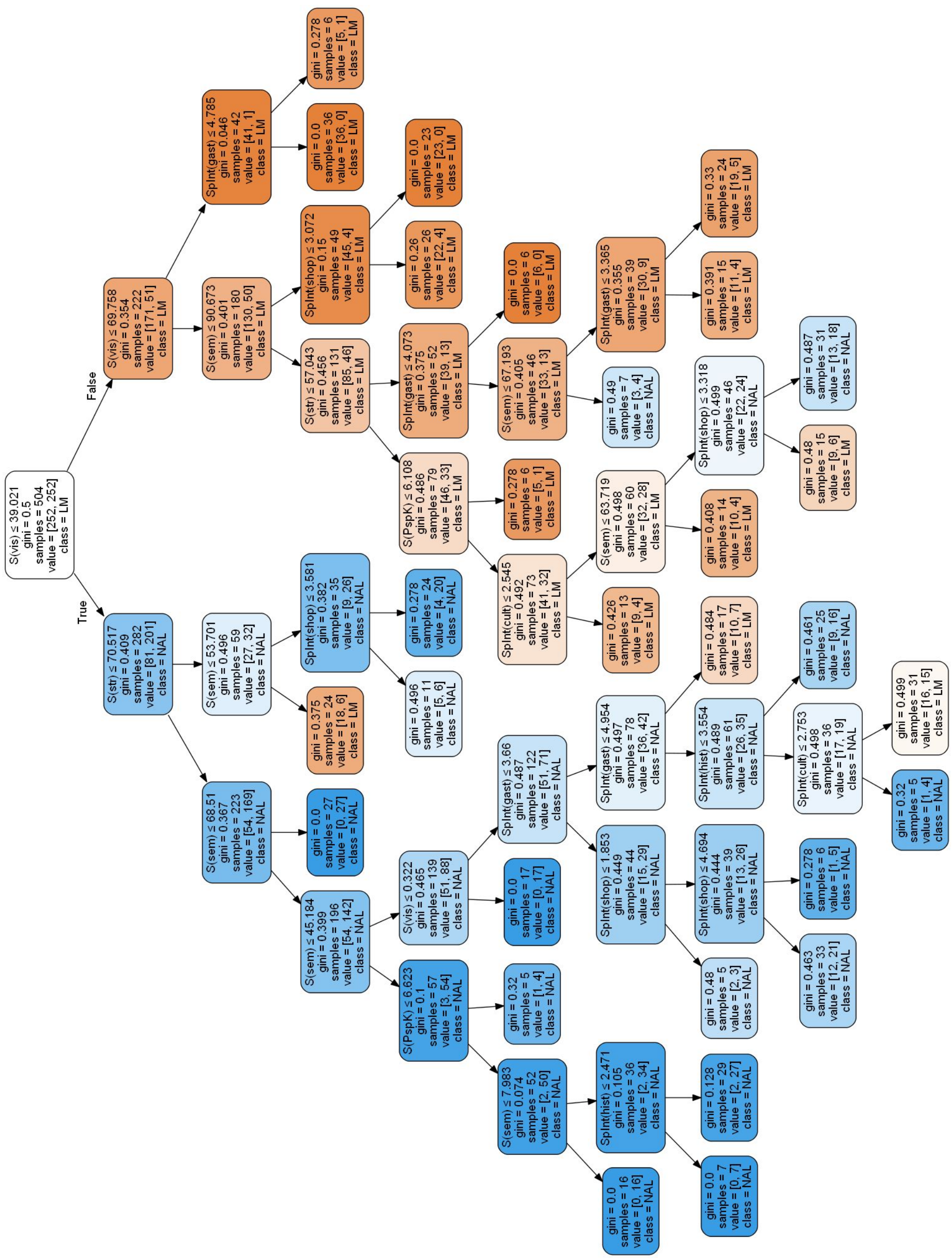


Figure 6.2: Personalised Decision Tree Model.

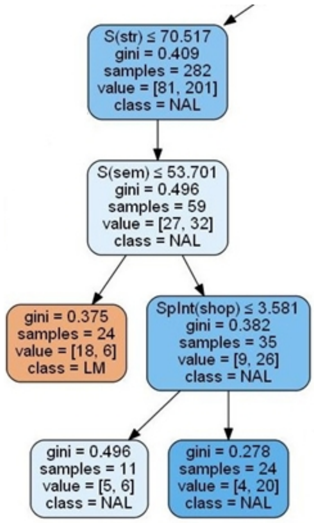


Figure 6.3: Part of the Personalised Decision Tree Model.

The model parameters (Table 6.8) define where decision tree growing stops before yielding all pure leaf nodes. In the case of a fully grown tree, there would be, for example, another decision after the evaluation of $s_{pInt(cult)}$ (at the bottom of Figure 6.2). Since this is not the case the terminal node of the LM-class shows $gini = 0.499$ meaning a 50% chance of classifying the objects correctly. In case of the NAL-class the gini-index is only 0.32 being more clear about the classification.

The PdTm is built on a training set. In case the tree would be fully grown it would likely overfit the training set and this might result in a low recall for the test set. We stop growing the tree before yielding all pure leave nodes, although this would mean a lower gini-index.

Figure 6.2 shows that the PdTm generates terminal nodes with the same class in a number of parts of the PdTm (e.g. class NAL in Figure 6.3). Why the algorithm does not stop one step earlier is because of the way the algorithm works. In case $min\_samples\_split$, $min\_samples\_leaf$, or max_depth is not reached the algorithm continues until it produces only leaf nodes that contain the minimum number of samples. As we set $min\_samples\_leaf = 5$ and $min\_samples\_split = 34$ as the optimal parameters for the optimal tree (Table 6.8), the algorithm stops before it can yield all pure leaf nodes. In Figure 6.3 the node on the right shows 4 samples of the class LM and 20 samples of the class NAL. Would the decision tree growing stop earlier it would produce the node above with 26 samples belonging to the class NAL and 9 samples identified as class LM which is far less useful.

We obtain a training accuracy of 78.17% and a slightly higher recall (78.97%,

Table 6.9: Recalls of the personalised models obtained with the test set.

	PwSm	PwPm	PdFc	PdTm
Recall [%]	40.95	32.33	35.34	66.38

Table 6.5). The difference is not as large as the difference obtained for the CdTm obtained on the training set (Table 5.8), meaning that the PdTm identifies LMs just as good as NALs. Section 6.2.4 investigates the performance of the trained PdTm on the test set.

6.2 Testing of the Models

We run the models on the test set and compare their identified landmarks with landmarks selected by survey participants. Then, we count identified landmarks of the models and calculate the recall (Table 6.9). We compare the results of the personalised models with a subsequent McNemar's test (Section 3.4). We always compare two models at a time with the test. In case there is no association between the two models, we expect the number of landmarks which the first model identifies but the second model does not to be equal to the number of landmarks which the second model identifies but the the first model does not. In this way we identify landmarks which change from identified correctly to unidentified and the other way round. Table 6.10 shows the results which we discuss in the following sections.

6.2.1 Personalised Weighted Sum Model

We apply the PwSm to the test set using Formula 6.3. The PwSm identifies 95 out of 232 landmarks correctly. It achieves a recall of 40.95% on the test set (Table 6.9). Subsequently, we perform a McNemar's test, comparing two models at a time (Table 6.10). The largest difference occurs between the PwSm and the PdTm with a McNemar's test statistic of 33.307 and a $p - value < 0.0001$. There are 101 discordant pairs. 21 pairs where a correctly identified landmark of the PwSm changes to an unidentified landmark of the PdTm. In contrast, there are 80 pairs where PdTm identifies a landmark but the PwSm does not. This difference is extremely statistically significant by conventional criteria ($p < 0.05$). The difference to the PwPm is considered to be extremely statistically significant. The test statistic is with 18.050 lower than the one to the PdTm. This is because between the PwSm and the PwPm no landmark changed from unidentified to correctly identified. The difference to the PdFc is considered to be statistically significant. There are only eight landmarks changing from unidentified to correctly identified and 21 landmarks changing the other way round.

6.2.2 Personalised Weighted Product Model

We identify landmarks with the PwPm applying the Formula 6.4. It identifies 75 out of 232 landmarks of the test set correctly. The performance of the PwPm is with 32.33% worse than those of the other models (Table 6.9). A subsequent McNemar's test shows that the largest difference is observed to the PdTm with 91 discordant pairs (Table 6.10). Six pairs where the PwPm identifies a landmark but the PdTm does not, and 85 pairs where the PdTm identifies a landmark but the PwPm does not. The difference to the PwSm is extremely statistically significant but has only 20 discordant pairs. Thereof, none of the landmarks changed from an unidentified of the PwSm to a correctly identified landmark of the PwPm. The p-value for the difference between the PwPm and the PdFc equals 0.146. This difference is not statistically significant. Only 12 landmarks change from an unidentified to a correctly identified landmark and only five the other way around between the PwPm and the PdFc.

6.2.3 Personalised Decision Flow Chart

The PdFc identifies landmarks following the flow in Figure 6.1. It identifies with 35.34% a recall similar to the PwPm, identifying seven more landmarks correctly (82 out of 232, Table 6.9). The McNemar's test confirms that the difference between the PwPm and the PdFc is considered to be not statistically significant with a p-value of 0.146 (Table 6.10). There are only 17 discordant pairs, which is the smallest observed difference between two models. In contrast, the difference to the PwSm is considered to be statistically significant. Although the PwSm only results in a 5.61% higher recall than the PdFc. In total there are 29 discordant pairs between these two models. The largest difference is observed to the result of the PdTm with a $p - value < 0.0001$ and a McNemar's test statistic of 54.793, which is considered to be extremely statistically significant. This means, there are only 10 pairs changing from an unidentified landmark of the PdTm to a correctly identified landmark of the PdFc whereas there are 82 pairs changing vice versa.

6.2.4 Personalised Decision Tree Model

We identify landmarks with the PdTm shown in Figure 6.2. The PdTm achieves with 66.38% the best result on the test set (Table 6.9). It identifies 154 landmarks of 232

Table 6.10: Results of McNemar's test of personalised models.

	PwPm	PdFc	PdTm
PwSm			
Unidentified → Identified	0	8	80
Identified → Unidentified	20	21	21
Test Statistic	18.050	4.966	33.307
p-value	<0.0001	0.026	<0.0001
PwPm			
Unidentified → Identified		12	85
Identified → Unidentified		5	6
Test Statistic		2.118	66.857
p-value		0.146	<0.0001
PdFc			
Unidentified → Identified			82
Identified → Unidentified			10
Test Statistic			54.793
p-value			<0.0001

correct. McNemar's test reveals an extremely statistically significant difference to all the other models ($p - value < 0.0001$) (Table 6.10). The largest difference occurs to the PwPm with a McNemar's test statistic of 66.857. There are 91 discordant pairs, thereof 85 changed from unidentified landmarks of the PwPm to correctly identified landmarks of the PdTm. The difference to the PdFc ranks second with a test statistic of 54.793. Here, 82 correctly identified landmarks of the PdTm changed to unidentified landmarks with the PdFc. Ten pairs are found where the PdFc identifies a landmark but the PdTm does not. The difference to the PwSm is as well extremely statistically significant with a McNemar's test statistic of 33.307. There are 101 discordant pairs thereof 80 changed from an unidentified landmark with the PwSm to a correctly identified landmark with the PdTm.

6.3 Results of the Training and the Testing Discussed

This section discusses the results of the training and the testing of the personalised landmark identification models and the achieved recalls. The recalls obtained on the training set (Table 6.5) vary for the PwSm, the PwPm, and the PdFc around 60%. The recall for the the PdTm is higher with 78.97% .

The PwSm as well as the PwPm obtain the best average recall with more than one combination of model parameters. We use the combination with the minimum whole numbers obtaining the best average recall for the PwSm and the PwPm respectively. The PwSm identifies the minimum combination with $p_{vis} = 2$, whereas the PwPm sets $p_{sem} = 2$. The other model parameters are set to one. An interesting finding is, that the PwPm obtains the same recall for $p_{vis} = 2$, $p_{sem} = 1$, and $p_{str} = 1$ as well as for $p_{vis} = 1$, $p_{sem} = 2$, and $p_{str} = 1$. However, we use the combination of model parameters with the minimum whole numbers for testing.

We expected from the knowledge of related work (Section 2.3) that a $s_{PspK} \leq 3$ would result in a higher semantic salience, whereas a $s_{PspK} > 3$ would result in a higher visual and structural salience. We expected that these tendencies reflect in the models parameters. However, the values of $\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$ in Table A.3 do not allow any conclusions that *prior spatial knowledge* or *personal interests* respectively have an impact on salience.

The overall results of the personalised landmark identification models show that the recalls achieved on the training set (Table 6.5) are higher than on the test set (Table 6.9). We already made this observation for the conventional models. The models better fit the data in the training set than in the test set. However, there might be a number of additional reasons for this result and we discuss them in the following paragraphs.

Different data distributions of training, validation, and test set The training, the validation and the test set might have different distributions. We ensured that training and validation sets have equal class distributions by applying stratified folds. The training and test set might show different distributions because we divide our dataset in two sets that do not overlap spatially. However, we expect that our data in the training and the test set overlap instead of having completely different distributions. Nevertheless, the test recall is much lower than the training recall. This especially applies for the PwSm, the PwPm, and the PdFc which are not learned from data as the PdTm. A possible solution for this problem might be to take the whole dataset (including all 10 decision points) and randomly shuffle it. Then, we might split the resulting dataset into training and test set. However, with this solution the training and test set would not be spatially independent anymore and that is important for our use case because we aim to develop personalised landmark identification models suitable in different spatial environments.

The models based on theory versus the machine learning model The PdTm performs with much higher recall compared to the other models (66.38%, Table 6.9). One reason for the deficits in recall of the PwSm, the PwPm, and the PdFc might be that they are based on theoretical considerations, whereas the PdTm learns from the training set. We only use the 10 validation folds to get an estimate of the performance of the PwSm, the PwPm, and the PdFc (Figure 3.4). Thus, the training folds are never touched and, thus, do not reflect in the model results. The PdTm learns the tree from the training set in order to identify decisions and identify LMs and NALs. This seems to lead to a higher recall of the PdTm.

Overfitting of the models to the training set Another issue which might lead to a low test recall is the already mentioned overfitting of the PdTm to the training folds and of the PwSm, the PwPm, and the PdFc to the validation folds. However, we applied cross-validation to avoid this phenomenon as much as possible. There might be additional solutions such as feature selection (Section 9.3) to tackle this problem. Since we need to consider all the personal dimensions we use the total number of dimensions and perform no subset selection during model training.

Overlapping of salience values of LMs and NALs A further issue concerns the overlapping of s_{vis}, s_{sem}, and s_{str} of LMs and NALs. There are objects in our survey, which are selected as both LM and NALs. This results in identical values for s_{vis}, s_{sem}, and s_{str} appearing for LMs as well as for NALs in the training set. Thus, an object with a particular combination of s_{vis}, s_{sem}, and s_{str} classified in the training set as a NAL might appear in the test set as a LM. This might result in a number of unidentified landmarks and, thus, keep the recall low.

The most important influence on the identification of landmarks is still unknown One possible interpretation of our results is that we have not yet found the most important dimensions for the identification of personalised landmarks. The present results suggest that *prior spatial knowledge* and *personal interests* are not the major dimensions for the identification of personalised landmarks. We might be either missing additional landmark or personal dimensions respectively or there might be dimensions, which have not been detected yet. For a discussion of other dimensions compare Section 8.1.

Chapter 7

Analyses and Comparison of the Models

The main goal of this chapter is to find out whether a personalised landmark identification model incorporating *prior spatial knowledge* and *personal interests* identifies more landmarks selected by survey participants than a conventional, non-personalised landmark identification model. We start this chapter with a sensitivity analysis to identify whether the personal dimensions influence the outcomes of the personalised models (Section 7.1). Subsequently, we use McNemar's test for the statistical comparison of the results of both conventional and personalised models (Section 7.2). We close the chapter with a conclusion of the results of the sensitivity analysis and the comparison of the models with regard to our hypothesis (Section 7.3).

7.1 Sensitivity Analysis of the Personalised Models

In this section we determine whether different values of the dimensions affect the outputs of the personalised landmark identification models. This sensitivity analysis investigates one dimension at a time, e.g. the effect that changes in *personal interests* (pInt) or *prior spatial knowledge* (PspK) ratings have on the outputs of the models. We vary one dimension from its minimum value to its maximum value, while keeping the values of the other dimensions constant (Section 3.3). The landmark dimensions s_{vis} and s_{sem} have five values to vary (0, 25, 50, 75, 100), s_{str} has only three values (0, 50, 100). pInt have five values (from one (*no* interest) to five (*very high* interest)) and PspK has seven possible values to vary (Table 4.2). Subsequently, we calculate the sensitivity index (SI) (Formula 3.7). We use examples from our test set. In cases where no suitable data are available, we use appropriate examples.

Table 7.1: Example for sensitivity analysis of the PwSm to the landmark dimensions.

ID	s_{vis}	s_{sem}	s_{str}	$s_{iLM(shop)}$	$s_{iLM(cult)}$	$s_{iLM(hist)}$	$s_{iLM(gast)}$
1	0...100	50	50	1	1	1	1
2	50	50	50	1	1	1	1
3	50	50	50	1	1	1	1

7.1.1 Personalised Weighted Sum Model

In this section we perform a sensitivity analysis of the PwSm. We present and discuss the results.

Sensitivity Analysis of the PwSm and Results

The result of the PwSm is a salience measure which determines whether an object functions as a landmark or not (Formula 6.3). The SI gives information about the magnitude and the direction in which the salience measure changes depending on the input values of the dimensions.

As a first step, we investigate the sensitivity of the PwSm to the landmark dimensions. We use an example because there are no suitable data in our test set. We start to investigate the sensitivity of the model to s_{vis} using the values in Table 7.1. We set constant values to $s_{PspK} = s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)}$ $= 3$ and we vary s_{vis} of object 1 from 0 to 100 using steps of 25. Table 7.2 shows that the number of identified landmarks changes with s_{vis}. In case $s_{vis} < 50$ the PwSm identifies object 2 and 3 as landmarks. In the case of $s_{vis} = 50$, all three objects show the same salience measure because they have identical values for the landmark dimensions. Object 1 is the unique landmark in case $s_{vis} \geq 75$. We calculate a SI = 0.56 for object 1. As the other objects do not change their salience measures during the sensitivity analysis, their SI is zero. However, the PwSm reacts sensitively to s_{vis}.

The sensitivity analysis of s_{vis} is representative for the other landmark dimensions. The Appendix shows the results for s_{sem} (Table A.5) and s_{str} (Table A.6). Their SI is 0.47. This means that s_{vis} exerts the highest influence of the landmark dimensions on the output of the PwSm.

For the initial weights for the PwSm we analysed the objects selected as landmarks by survey participants with different pInt and PspK ratings. For each combination we determined the average of visual, semantic, and structural salience (Section 6.1.1). As described in Section 5.3, there are combinations of PspK and pInt ratings which

Table 7.2: Results of sensitivity analysis of PwSm to s_{vis}.

ID	s_{vis}					SI
	0	25	50	75	100	
1	1.56	2.06	**2.56**	**3.06**	**3.56**	0.56
2	**2.56**	**2.56**	**2.56**	2.56	2.56	0
3	**2.56**	**2.56**	**2.56**	2.56	2.56	0

do not appear in the training set. The result is that the average salience $\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$ are not available for all combinations (Table A.3). Thus, we are not able to calculate the salience measure with the PwSm for all combinations of PspK and pInt ratings. For the sensitivity analysis this means that if we vary the values of PspK and keep the values of the other dimensions constant, there are at most four average salience values available. Consider the pInt ratings $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$. There are average salience values for $s_{PspK} = \{1, 2, 3, 7\}$ available in Table A.3. Consequently, we are not able to calculate the salience measure of the PwSm for $s_{PspK} = \{4, 5, 6\}$.

Table 7.3: Objects at decision point 8 (Figure A.9).

ID	s_{vis}	s_{sem}	s_{str}	$s_{iLM(shop)}$	$s_{iLM(cult)}$	$s_{iLM(hist)}$	$s_{iLM(gast)}$
O_1	50	50	100	0	1	1	0
O_2	0	50	50	0	0	0	0
O_3	50	75	50	1	0	1	0
O_4	0	50	50	1	0	0	0
O_5	50	50	100	0	1	1	0

For the investigation of the model's sensitivity to s_{PspK} we use the objects at decision point 8 (Table 7.3). Table 7.4 shows the results for the available combinations of PspK and pInt ratings. s_{PspK} has a different effect on the salience measure of each object depending on s_{vis}, s_{sem}, and s_{str}. The SI varies in magnitude (SI $\in$ {-0.27 - -0.07}) giving information of the direction of the change of the salience measure. The salience measure decreases with an increasing value of s_{PspK}. This has an impact on the number of detected landmarks. The PwSm identifies two objects as landmarks in case $s_{PspK} \leq 3$ (Table 7.4, O_1 and O_5). In case $s_{PspK} = 7$, the PwSm identifies an additional object (O_3) as landmark. The average sensitivity index (avgSI) is -0.20.

Table 7.4: Results of sensitivity analysis of PwSm to s_{PspK}.

ID	s_{PspK}				SI
	1	2	3	7	
O_1	**3.80**	**3.50**	**3.06**	**3.00**	-0.27
O_2	1.80	1.63	1.39	1.50	-0.20
O_3	3.20	3.00	2.75	**3.00**	-0.07
O_4	1.80	1.63	1.39	1.50	-0.20
O_5	**3.80**	**3.50**	**3.06**	**3.00**	-0.27
avgSI					-0.20

We conclude that the dimension PspK affects the outputs of the PwSm since the number of objects identified as landmarks changes according to s_{PspK}.

Next, we evaluate the sensitivity of the PwSm to pInt. We take again the objects from decision point 8 as an example (Table 7.3). When varying pInt at most two average salience values with the same combinations of PspK and pInt ratings are available (Table A.3). We set constant values $s_{PspK} = 7$ and $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$ and vary $s_{pInt(shop)}$, $s_{pInt(cult)}$, and $s_{pInt(hist)}$ successively. We have to take different values for $s_{pInt(gast)}$ to be able to use an example from our test set. We set $s_{PspK} = 3$, $s_{pInt(shop)} = 2$, and $s_{pInt(cult)} = s_{pInt(hist)} = 3$ when investigating $s_{pInt(gast)}$. However, the sensitivity analysis reaches the same conclusion for $s_{pInt(gast)}$ as for the other pInt (Table 7.5 and Table 7.6). The PwSm identifies different numbers of landmarks according to the interest ratings. Generally, we can identify a difference between a pInt rating ≤ 3 and a rating > 3. Similarly to the sensitivity analysis results of s_{PspK} the PwSm identifies two landmarks for a pInt rating ≤ 3 and three landmarks for a rating > 3. The magnitude of the sensitivity varies and results in an avgSI $\in$ {-0.10 - 0.04}.

Results of the Sensitivity Analysis of the PwSm discussed

The PwSm shows the lowest sensitivity to variation in the inputs of the pInt, followed by the sensitivity to PspK. We found the highest sensitivity to the landmark dimensions, this means that they have the highest influence on the outcomes of the PwSm.

We investigate the model's sensitivity to the landmark dimensions with an example because of the lack of appropriate data. The influence of a change in a landmark dimension on the model's results is dependent on the values of the landmark dimensions of the other objects at an investigated decision point. For example, in case all other objects in our example would have $s_{vis} = s_{sem} = s_{str} = 100$ (Table A.7), object 1 would only become a landmark in case $s_{vis} = 100$ (Table A.8).

Table 7.5: Results of sensitivity analysis of PwSm to $s_{pInt(shop)}$ and $s_{pInt(cult)}$.

ID	$s_{pInt(shop)}$		SI	$s_{pInt(cult)}$		SI
	2	4		3	4	
O_1	**3.29**	**3.00**	-0.10	**3.33**	**3.00**	-0.11
O_2	1.57	1.50	-0.05	1.67	1.5	-0.11
O_3	3.00	**3.00**	0.00	3.17	**3.00**	-0.06
O_4	1.57	1.50	-0.05	1.67	1.5	-0.11
O_5	**3.29**	**3.00**	-0.10	**3.33**	**3.00**	-0.11
avgSI			-0.06			-0.10

Table 7.6: Results of sensitivity analysis of PwSm to $s_{pInt(hist)}$ and $s_{pInt(gast)}$.

ID	$s_{pInt(hist)}$		SI	$s_{pInt(gast)}$		SI
	3	4		2	4	
O_1	**3.15**	**3.00**	-0.05	**3.13**	**3**	-0.04
O_2	1.45	1.50	0.03	1.38	1.5	0.08
O_3	2.83	**3.00**	0.06	2.69	**3.00**	0.10
O_4	1.45	1.50	0.03	1.38	1.5	0.08
O_5	**3.15**	**3.00**	-0.05	**3.13**	**3.00**	-0.04
avgSI			0.01			0.04

The PwSm differentiates between $s_{PspK} \leq 3$ and $s_{PspK} = 7$. This makes sense because a rating ≤ 3 means that the survey participant is familiar with the street intersection, whereas a rating of seven means that the survey participant has never been there. Table A.3 shows that for $s_{PspK} = 7$ and $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$ average salience values of $\overline{s}_{sem} = 100$ and $\overline{s}_{vis} = \overline{s}_{str} = 50$ apply. This results in a relative weight $w_{semRel} = 2$ (Formula 6.1) which is higher than for the other PspK ratings. The result is that object O_3 becomes a landmark for $s_{PspK} = 7$. This is in contrast to the findings of the literary research (Section 2.3) saying that for unfamiliar people visual salience is more important than semantic salience. However, since the combination of PspK and pInt rating (s_{PspK}

$= 7$ and $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$) appears only once in the training set (Table A.3, Column *No*), this result might not accurately reflect the actual preferences of travellers with these ratings. The other PspK ratings in Table 7.4 with $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$ are chosen more often by the survey participants, and, therefore, seem to be reliable.

Although we use slightly different constant values to show sensitivity to all pInt an interest rating ≤ 3 results in two, wheres a high interest rating > 3 results in three identified landmarks. Thus, the output of the PwSm changes according to the ratings of pInt. We identify for object O_3 in the sensitivity analysis of $s_{pInt(shop)}$ a salience measure of 3.00 in any case (Table 7.5). This results in a SI = 0. While we do not identify this object for $s_{pInt(shop)} = 2$, for $s_{pInt(shop)} = 4$ the object becomes a landmark although it does not change its salience measure. However, because the other objects change their salience measures object O_3 becomes one of the most salient ones. The highest magnitude of SI shows $s_{pInt(cult)}$. Nearly for all objects the SI is -0.11, with a decreasing salience measure.

7.1.2 Personalised Weighted Product Model

The sensitivity analysis of the PwPm is similar to the analysis of the PwSm. In the following we present the results and discuss them.

Sensitivity Analysis of the PwPm and Results

We use Formula 6.4 to calculate the salience measure that defines which object becomes a landmark. We calculate the SI giving us information about the magnitude of the differences of the salience measures while changing input values of the dimensions.

We start with a sensitivity analysis of the PwPm to the landmark dimensions. We use the same example as for the sensitivity analysis of the PwSm (s_{vis} and the values in Table 7.1). We set constant values to $s_{PspK} = s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 3$. We vary s_{vis} of object 1 from 0 to 100 using steps of 25. Table 7.7 shows that the number of identified landmarks changes with s_{vis}. It is identical to the results of the PwSm: in case $s_{vis} < 50$ the PwPm identifies object 2 and 3 as landmarks, in case $s_{vis} = 50$ all three objects are identified as landmarks, and in case $s_{vis} \geq 75$ object 1 becomes the unique landmark. The SI is 1 for object 1 because the value for D_{min} is zero (Formula 3.7). We can conclude that the PwPm reacts sensitively to s_{vis}. The sensitivity analysis of s_{vis} is representative for the other landmark dimensions s_{sem} and s_{str}. The Appendix shows the results for s_{sem} (Table A.9) and s_{str} (Table A.10). As $s_{vis} = 0$ leads to a $D_{min} = 0$, SI = 1 applies

Table 7.7: Results of sensitivity analysis of PwPm to s_{vis}.

s_{vis}	ID		
	1	2	3
0	0	**42412775.26**	**42412775.26**
25	21206387.63	**42412775.26**	**42412775.26**
50	**42412775.26**	**42412775.26**	**42412775.26**
75	**63619162.89**	42412775.26	42412775.26
100	**84825550.52**	42412775.26	42412775.26
SI	1	0	0

in any case. The SI is 1 for all three landmark dimensions, and consequently, they have the same influence on the model results.

Here, as in the case of the sensitivity analysis of the PwSm, we have to handle combinations of PspK and pInt ratings not appearing in the test set. For the sensitivity analysis this means that for varying PspK and constant pInt, there are at most four average salience values available. For the investigation of the PwPm to PspK, we refer again to the objects of decision point 8 (Table 7.3). We set constant values to $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$. We vary s_{PspK} from one 1 to 7. Table 7.8 shows the results for the values of $s_{PspK} = \{1, 2, 3, 7\}$. The other combinations of PspK and pInt ratings are not available.

Table 7.8: Results of sensitivity analysis of PwPm to s_{PspK}.

ID	s_{PspK}				SI
	1	2	3	7	
O_1	**1366702592.41**	197642353.76	18783314.88	312500000	-3.37
O_2	0	0	0	0	0
O_3	1250564384.05	**198313053.74**	**24049552.14**	**791015625**	-0.58
O_4	0	0	0	0	0
O_5	**1366702592.41**	197642353.76	18783314.88	312500000	-3.37
avgSI					-1.47

s_{PspK} has different effects on the salience measures each object has depending on the values of s_{vis}, s_{sem}, and s_{str}. In case one of the landmark dimensions is equal to zero the salience measure is also zero. In this case the SI is not calculable because of

Table 7.9: Results of sensitivity analysis of PwPm to $s_{pInt(shop)}$.

ID	$s_{pInt(shop)}$			SI
	1	2	4	
O_1	39434834.03	**625000000**	18783314.88	-1.10
O_2	0	0	0	0
O_3	**50432651.36**	527343750	**24049552.14**	-1.10
O_4	0	0	0	0
O_5	39434834.03	**625000000**	18783314.88	-1.10
avgSI				-0.66

a division by zero (Formula 3.7). To be able to average the SI, we treat these values as zero. This results in a SI $\in$ {-3.37 - 0} for s_{PspK}. The PspK has an effect on the identified landmark as well as on the number of identified landmarks. In case s_{PspK} = 1 the PwPm identifies two objects as landmarks (O_1 and O_5). For all the other investigated PspK ratings the PwPm identifies only one landmark which is different from the one identified with $s_{PspK} = 1$. For the objects O_1 and O_5 we calculate a high SI = -3.37. These are the objects identified as a landmark for $s_{PspK} = 1$ but not for the other ratings. Consequently, the PwPm is sensitive to the PspK of a traveller.

In the next step we evaluate the sensitivity of the PwPm for the pInt. We use again decision point 8 as an example (Table 7.3). At most three average salience values with the same combination of PspK and pInt ratings are available in our test set (Table A.3). We set constant values to $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 4$ and vary the values of one pInt at a time. We have to set different values for the constant value of s_{PspK} to be able to demonstrate sensitivity with data from our test set. Thus, for $s_{pInt(shop)}$ and $s_{pInt(cult)}$: $s_{PspK} = 3$ and for $s_{pInt(hist)}$ and $s_{pInt(gast)}$: $s_{PspK} = 1$. We have to set different constant values for $s_{pInt(gast)}$ in case we want to use an example from our test set. We set $s_{PspK} = 1$, $s_{pInt(shop)} = 5$, and $s_{pInt(cult)} = s_{pInt(hist)} = 3$.

Table 7.9 - Table 7.12 show the results of the sensitivity analyses of the pInt. The sensitivity analysis of $s_{pInt(shop)}$ considers three stages of pInt: $s_{pInt(shop)} \in \{1, 2, 4\}$ (Table 7.9). It reveals a sensitivity resulting in either one or two identified landmarks for different values of $s_{pInt(shop)}$. Although the salience measure increases between an interest rating of $s_{pInt(shop)} = 1$ and $s_{pInt(shop)} = 2$ the SI is negative. This shows that the salience measure of the PwPm decreases between the minimum rating ($s_{pInt(shop)}$ = 1) and the maximum rating ($s_{pInt(shop)} = 4$).

Table 7.10: Results of sensitivity analysis of PwPm to $s_{pInt(cult)}$.

ID	$s_{pInt(cult)}$		SI
	3	4	
O_1	**36550221.73**	18783314.88	-0.95
O_2	0	0	0
O_3	33942986.18	**24049552.14**	-0.41
O_4	0	0	0
O_5	**36550221.73**	18783314.88	-0.95
avgSI			-0.46

Table 7.11: Results of sensitivity analysis of PwPm to $s_{pInt(hist)}$.

ID	$s_{pInt(hist)}$		SI
	3	4	
O_1	7179364.72	**1366702592.41**	0.99
O_2	0	0	0
O_3	**9725356.88**	1250564384.05	0.99
O_4	0	0	0
O_5	7179364.72	**1366702592.41**	0.99
avgSI			0.60

Table 7.12: Results of sensitivity analysis of PwPm to $s_{pInt(gast)}$.

ID	$s_{pInt(gast)}$		SI
	3	4	
O_1	**965282565.59**	52265689.43	-17.47
O_2	0	0	0
O_3	851988913.52	**57366173.64**	-13.85
O_4	0	0	0
O_5	**965282565.59**	52265689.43	-17.47
avgSI			-9.76

The sensitivity analyses of $s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)}$ reach the same conclusion: they identify different numbers of landmarks for $s_{pInt} = 3$ and $s_{pInt} = 4$ (Table 7.10 - Table 7.12). However, the magnitude of the sensitivity varies considerably and results in an avgSI $\in$ {-9.76 - 0.60}. The negative SI shows that the salience measure of the PwPm decreases with and increasing pInt rating.

Results of the Sensitivity Analysis of the PwPm discussed

The PwPm shows sensitivity to all dimensions - landmark as well as personal dimensions. We observe the lowest avgSI for $s_{pInt(shop)}$, $s_{pInt(cult)}$ and $s_{pInt(hist)}$, whereas $s_{pInt(gast)}$ shows the highest observed SI. The SI for s_{PspK} and the SIs for the landmark dimensions range inbetween.

Similar to the PwSm we investigate the model's sensitivity to the landmark dimensions with an example. It shows that the influence of a change in a landmark dimension results in a change of the identified or the number of identified landmarks respectively. It depends on the values of the other objects at an investigated decision point, whether an object becomes a landmark.

The PwPm reacts sensitively to the inputs of PspK. In case $s_{PspK} = 1$ two landmarks are identified which are different from the one identified for the other PspK ratings. However, the result of the sensitivity analysis of PspK is not as obvious as for the PwSm. We are not able to confirm a differentiation between no *prior spatial knowledge* and *prior spatial knowledge*. There are only indications that the PwPm differentiates between a traveller familiar with the street intersection and the area and all the other possible stages of PspK.

Considering the sensitivity analysis of the PwPm concerning the pInt, two average salience values of PspK and pInt ratings are available for a variation of $s_{pInt(cult)}$, $s_{pInt(hist)}$, and $s_{pInt(gast)}$. The sensitivity analysis shows that a $s_{pInt} = 3$ (*medium*) results in different and, in addition, in a different number of identified landmarks than a $s_{pInt} = 4$ (*high*). However, for $s_{pInt(cult)} = s_{pInt(gast)} = 3$ the PwPm identifies two, whereas for $s_{pInt(hist)} = 3$ the PwPm identifies only one landmark (Tables 7.10 - 7.12). We can conclude that the PwPm reacts sensitively to the inputs of the pInt $s_{pInt(cult)}$, $s_{pInt(hist)}$, and $s_{pInt(gast)}$. For $s_{pInt(shop)}$ we also detect sensitivity, but are not able to differentiate between *interested* and *not interested* participants. This means the PwPm identifies the same landmark for a $s_{pInt(shop)} = 1$ (*no*) and $s_{pInt(shop)} = 4$ (*high*). The SI varies considerably between the pInt, ranging from an avgSI = -9.76 for $s_{pInt(gast)}$ to an avgSI = 0.6 for $s_{pInt(hist)}$. The highest influence of the pInt on the outcomes of the PwPm shows $s_{pInt(gast)}$.

Table 7.13: Results of sensitivity analysis of PdFc to s_{vis} with $s_{PspK} = 7$.

ID	s_{vis}					SI
	0	25	50	75	100	
1	NAL	NAL	LM	LM	LM	1
2	LM	LM	LM	NAL	NAL	1
3	LM	LM	LM	NAL	NAL	1
avgSI						1

Table 7.14: Results of sensitivity analysis of PdFc to s_{vis} with $s_{PspK} = 1$.

ID	s_{vis}					SI
	0	25	50	75	100	
1	LM	LM	LM	LM	LM	0
2	LM	LM	LM	LM	LM	0
3	LM	LM	LM	LM	LM	0
avgSI					0	

7.1.3 Personalised Decision Flow Chart

This section investigates whether the PdFc reacts sensitively to the inputs of both the landmark as well as the personal dimensions. Furthermore, it presents and discusses the results.

Sensitivity Analysis of the PdFc and Results

We follow the flow in Figure 6.1 which results in one or more identified landmarks for a decision point. We investigate the result of the model when the dimension is set to its maximum and to its minimum respectively. The SI $= 1$ in case an investigated object changes from identified LM to NAL or vice versa.

As a first step, we evaluate the sensitivity of the PdFc to the landmark dimensions. We set the personal dimensions to the constant values $s_{PspK} = 7$ and $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = s_{pInt(gast)} = 1$. Consider the example in Table 7.1. We vary s_{vis} of object 1 from 0 to 100. Table 7.13 shows that the identified landmarks change with s_{vis}. In case $s_{vis} \leq 25$ the PdFc identifies the objects 2 and 3 as landmarks. In case $s_{vis} = 50$ all objects of the example are identified as landmark and in case $s_{vis} \geq 75$ object 1 is the only object qualifying as a landmark. Thus, the PdFc reacts

Table 7.15: Objects at decision point 7 (Figure A.8).

ID	s_{vis}	s_{sem}	s_{str}	$s_{iLM(shop)}$	$s_{iLM(cult)}$	$s_{iLM(hist)}$	$s_{iLM(gast)}$
O_1	25	75	100	1	1	0	1
O_2	25	50	100	0	0	0	1
O_3	50	25	50	0	0	1	0
O_4	25	50	50	1	0	0	0

sensitively to the inputs of s_{vis}.

Now we set $s_{PspK} = 1$ keeping the values of the pInt as they are. In this case the PdFc identifies all objects as landmarks because the model only examines s_{sem} (Figure 6.1). Thus, the sensitivity gets lost and SI = 0 (Table 7.14).

The PdFc behaves vice versa for s_{sem}. As long as $s_{PspK} \leq 3$ the PdFc is sensitive to the inputs of s_{sem}. s_{str} does not appear in the model and, therefore, we cannot say anything about the sensitivity of s_{str}.

The second step investigates the sensitivity of the PdFc to the personal dimensions. Unfortunately, we are not able to demonstrate the model's sensitivity to all dimensions with the data from our test set, thus, for some dimensions we have to refer to examples.

For s_{PspK} we are able to demonstrate the model's sensitivity using decision point 7 and its objects (Table 7.15) . We take constant values $s_{pInt(shop)} = 2$, $s_{pInt(cult)} = s_{pInt(hist)} = 3$, and $s_{pInt(gast)} = 4$ from our test set. We vary s_{PspK} from 1 to 7. Table 7.16 shows the results. There are no values for $s_{PspK} = \{2, 4, 5\}$ and the defined constant values for the pInt in our test set. Nevertheless, there are enough values for s_{PspK} to draw a conclusion about the model's sensitivity. We can see in the Table 7.16 that in case $s_{PspK} \leq 3$, object O_1 is identified as landmark, in case $s_{PspK} \geq 6$, it changes to object O_3. For the objects O_1 and O_3 at decision point 7 the SI = 1 and the avgSI = 0.5. We conclude that the PdFc is sensitive to s_{PspK}.

We are not able to demonstrate sensitivity to pInt with data from our test set. We take the objects at decision point 5 and exemplary personal data as an example (Table 7.17). The PdFc differentiates only between *interested* and *not interested*. It follows $s_{pInt} \in \{0, 1\}$. We start with evaluating the sensitivity of the PdFc to $s_{pInt(gast)}$. We use exemplary constant values $s_{PspK} = 1$ and $s_{pInt(shop)} = s_{pInt(cult)} = s_{pInt(hist)} = 0$. Table 7.18 shows the result for sensitivity analysis of $s_{pInt(gast)}$. In case $s_{pInt(gast)} = 0$, the PdFc identifies two landmarks. Given $s_{pInt(gast)} = 1$, the number of landmarks decreases to one. The SI for object O_2 is 1. The avgSI = 0.25 for decision point 5. Thus, the PdFc reacts sensitively to the inputs of $s_{pInt(gast)}$.

Table 7.16: Results of sensitivity analysis of PdFc to s_{PspK}.

ID	s_{PspK}				SI
	1	3	6	7	
O_1	LM	LM	NAL	NAL	1
O_2	NAL	NAL	NAL	NAL	0
O_3	NAL	NAL	LM	LM	1
O_4	NAL	NAL	NAL	NAL	0
avgSI					0.5

Table 7.17: Objects at decision point 5 (Figure A.6).

ID	s_{vis}	s_{sem}	s_{str}	$s_{iLM(shop)}$	$s_{iLM(cult)}$	$s_{iLM(hist)}$	$s_{iLM(gast)}$
O_1	50	75	100	1	0	1	1
O_2	50	75	100	1	0	1	0
O_3	25	50	100	0	0	0	0
O_4	25	50	50	1	0	0	1

The other pInts show the same behaviour.

Now we set $s_{pInt(shop)} = 1$. The other constant values remain the same. As a result the PdFc identifies the same two landmarks (Table 7.19) independent of the value of $s_{pInt(gast)}$. The reason for this is that objects O_1 and O_2 both belong to the topic of interest *shopping*. This means, the PdFc identifies these two objects as landmarks as long as $s_{pInt(shop)} = 1$. This results in a avgSI = 0. In this case the PdFc is no longer sensitive to $s_{pInt(gast)}$. The other pInt show the same behaviour.

Results of the Sensitivity Analysis of the PdFc discussed

We identify the sensitivity of the PdFc to landmark as well as to personal dimensions. However, there are some restrictions. The model's sensitivity to s_{vis} and s_{sem} respectively is highly dependent on s_{PspK}. This is obvious because the model makes a clear distinction between $s_{PspK} \leq 3$ (meaning *prior spatial knowledge*) and $s_{PspK} > 3$ (meaning no *prior spatial knowledge*) (Figure 6.1). In case $s_{PspK} > 3$ the model investigates only s_{vis} while s_{sem} is not considered. For $s_{PspK} \leq 3$ it is just the opposite.

The PdFc makes a distinction between $s_{PspK} \leq 3$ and $s_{PspK} \geq 6$. Remember, survey participants did not choose the $s_{PspK} = 4$ or $s_{PspK} = 5$, therefore, we are

Table 7.18: Results of sensitivity analysis of PdFc to $s_{pInt(gast)}$.

ID	$s_{pInt(gast)}$		SI
	0	1	
O_1	LM	LM	0
O_2	LM	NAL	1
O_3	NAL	NAL	0
O_4	NAL	NAL	0
avgSI			0.25

Table 7.19: Results of sensitivity analysis of PdFc to $s_{pInt(gast)}$ with $s_{pInt(shop)} = 1$.

ID	$s_{pInt(gast)}$		SI
	0	1	
O_1	LM	LM	0
O_2	LM	LM	0
O_3	NAL	NAL	0
O_4	NAL	NAL	0
avgSI			0

not able to draw conclusions about theses values. Nevertheless, we may conclude that the PdFc divides between familiarity and no familiarity and dependent on that different landmarks are identified and we expected that because of the first decision of the PdFc (Figure 6.1).

Table 7.18 shows sensitivity to pInt - in this case using the example of $s_{pInt(gast)}$. However, all pInt behave the same way. This is due to the flow of the PdFc. The PdFc has a number of objects available after passing the process for the max(s_{vis}) or max(s_{sem}) respectively. The objects satisfying the decision $s_{iLM} = s_{pInt} = 1$ are identified as landmarks (Figure 6.1). This means, in most cases this decision narrows down the number of objects available, and, thus, the number of identified landmarks. The PdFc reacts sensitively to the inputs of the pInt under investigation in case all the other $s_{pInt} = 0$. However, the sensitivity to the pInt under investigation gets lost as soon as the analysis is extended to other pInt when there are objects that are part of the interest. In this case the prerequisite $s_{iLM} = s_{pInt} = 1$ is fulfilled for more than one object and, thus, the results of the PdFc are more identified landmarks.

Table 7.20: Results of sensitivity analysis of PdTm to landmark dimensions (B = depends on other dimensions whether LM or NAL).

Landmark	0	25	50	75	100	SI
s_{vis}	B	B	B	LM	LM	1
s_{sem}	B	B	B	B	B	1
s_{str}	B		B		B	1

7.1.4 Personalised Decision Tree Model

In this section we evaluate the PdTm with a sensitivity analysis. We present and discuss the results.

Sensitivity Analysis of the PdTm and Results

The result of the PdTm is not a salience measure (as in the case of the PwSm or the PwPm) but a class. The results are either LM or NAL. This results in a SI $\in$ (0, 1). When considering the design of the PdTm, it is striking that the dimensions influence each other (Figure 6.2). This means, the sensitivity of the PdTm to a dimension is dependent on the values of the other dimensions. For example, the branch where $s_{pInt(shop)}$ is located is only entered when $39.021 < s_{vis} \leq 69.758$.

First, we start evaluating the sensitivity of the PdTm to the landmark dimensions. It turns out that the PdTm is sensitive to all three dimensions (Table 7.20). This is particularly obvious for s_{vis} because in case $s_{vis} > 69.758$ the object is a landmark in any case. For s_{sem} and s_{str} it depends on the values of the other dimensions, whether the PdTm reacts sensitively. For example, for $s_{vis} \leq 39.021$ and $s_{str} > 70.517$ it depends exclusively on s_{sem} whether an object becomes a LM or a NAL (Figure 6.2).

Next, we evaluate the sensitivity of the PdTm to the personal dimensions. The right branch of the tree hosts s_{PspK}. However, to enter the branch a number of requirements must be met: $39.021 < s_{vis} \leq 69.758$, $s_{sem} \leq 90.673$ and $s_{str} \leq 57.043$. In case $s_{PspK} > 6.108$ an object is identified as a landmark in any case. In case $s_{PspK} \leq 6.108$ the classification as LM or NAL depends on $s_{pInt(cult)}$, s_{sem}, and $s_{pInt(shop)}$. In case $s_{pInt(cult)} > 2.545$, $s_{sem} > 63.719$, and $s_{pInt(shop)} > 3.318$ the PdTm identifies an object as a NAL. Thus, as soon as $s_{pInt(shop)} \leq 3.318$ the PdTm identifies an object as a landmark. This means, the model shows a clear sensitivity to the input values of $s_{pInt(shop)}$ (Table 7.21).

$s_{pInt(cult)}$, $s_{pInt(hist)}$, and $s_{pInt(gast)}$ are all located in the left branch of the tree

Table 7.21: Results of sensitivity analysis of PdTm to personal dimensions (B = depends on other dimensions whether LM or NAL).

Personal	1	2	3	4	5	6	7	SI
s_{PspK}	B	B	B	B	B	B	LM	1
$s_{pInt(shop)}$	LM	LM	LM	NAL	NAL			1
$s_{pInt(cult)}$	NAL	NAL	LM	LM	LM			1
$s_{pInt(hist)}$	B	B	B	NAL	NAL			1
$s_{pInt(gast)}$	B	B	B	B	LM			1

(Figure 7.1). We use as constant values $s_{vis} = 25$ and $s_{sem} = s_{str} = 50$ to enter this branch. We do not need to set values for s_{PspK} and $s_{pInt(shop)}$ because they do not appear in this part of the tree. We set $s_{pInt(cult)} = 2$, $s_{pInt(hist)} = 3$, and $s_{pInt(gast)} = 4$. We vary the values of the analysed dimensions ($s_{pInt(cult)}$, $s_{pInt(hist)}$, and $s_{pInt(gast)}$) from their minimum value (1) to their maximum value (5) and investigate the resulting SI (Table 7.21).

The sensitivity analysis shows that the PdTm reacts sensitively to all three dimensions, but it is for $s_{pInt(hist)}$ and $s_{pInt(gast)}$ dependent on the values of the other dimensions. First, we investigate $s_{pInt(cult)}$. The PdTm is sensitive to $s_{pInt(cult)}$ with the above defined constant values. As soon as $s_{pInt(cult)} > 2.753$ the PdTm identifies an object as a landmark (Figure 7.1).

Second, we investigate $s_{pInt(hist)}$. The model's sensitivity to $s_{pInt(hist)}$ is partly dependent on the values of $s_{pInt(cult)}$. In case $s_{pInt(hist)} > 3.554$ an object is a NAL in any case. In case $s_{pInt(hist)} \leq 3.554$ it depends on $s_{pInt(cult)}$ whether an object is identified as a LM or a NAL.

The PdTm is sensitive to $s_{pInt(gast)}$ (Table 7.21). In case $s_{pInt(gast)} > 4.954$ an object becomes a landmark. The model's sensitivity to $s_{pInt(gast)}$ is also dependent on $s_{pInt(cult)}$ and $s_{pInt(hist)}$. As soon as $s_{pInt(hist)} > 3.554$, the object becomes a NAL. In case $s_{pInt(hist)} \leq 3.554$ and $s_{pInt(cult)} > 2.753$, the object becomes a LM. $s_{pInt(gast)}$ appears also in the right branch of the decision tree confirming the threshold between a rating of 4 (*high* interest) and a rating of 5 (*very high* interest) ($s_{pInt(gast)} > 4.073$). In the right branch of the PdTm it is dependent on s_{sem}, whether an object becomes a NAL or a LM in case $s_{pInt(gast)} \leq 4.073$. In case s_{sem} remains *50* (as defined above for constant values), the object becomes a NAL.

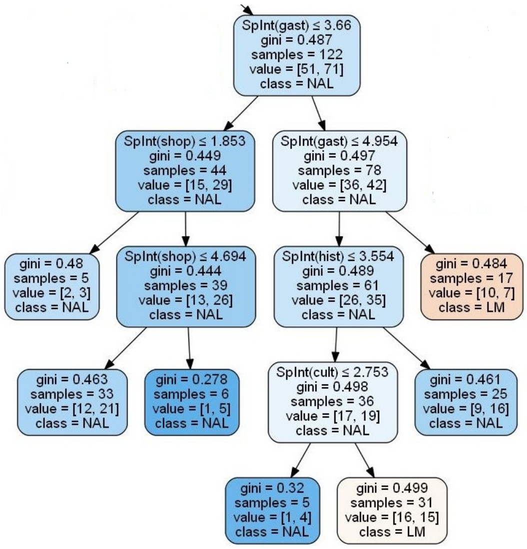


Figure 7.1: Decision Branch of the Personalised Decision Tree Model.

Results of the Sensitivity Analysis of the PdTm discussed

The PdTm shows sensitivity to all dimensions - landmark and personal respectively. For $s_{vis} \geq 75$ an object is a landmark in any case, whereas the sensitivity of the PdTm to the other landmark dimensions depends on correlations with the other dimensions.

The PdTm divides between $s_{PspK} = 7$ (no *prior spatial knowledge*) and all the other *prior spatial knowledge* ratings. Survey participants did not choose the ratings 4 and 5. The results show that on average only one survey participant chose $s_{PspK} = 6$ (Table 5.3). This suggests that this rating does not influence the splitting during tree growing. Thus, the distinction between no familiarity at all and the other ratings seems to be plausible. For the pInt we observe that $s_{pInt(shop)}$, $s_{pInt(cult)}$, and $s_{pInt(hist)}$ either split between $s_{pInt} = 2$ and $s_{pInt} = 3$ or $s_{pInt} = 3$ and $s_{pInt} = 4$. This means, the tree detects a difference between survey participants which are interested and which are not. $s_{pInt(gast)}$ represents an exception of the pInt as it makes the distinction between *very high* and all the other ratings.

There are personal dimensions appearing on more than one leave in the tree. However, most of them appear with the same decisions. Somehow different, however, behaves $s_{pInt(cult)}$. It appears twice - once with the decision $s_{pInt(cult)} \leq 2.753$ and once again with $s_{pInt(cult)} \leq 2.545$. Although the threshold values are similar, the decision whether the object is a LM or a NAL is contradictory. On the basis of the PdTm it is nevertheless comprehensible because whether an object classifies as a landmark is also dependent on the values of the other dimensions.

7.1.5 Results of the Sensitivity Analyses

In this chapter we perform a sensitivity analysis of our models to investigate the relationship between the input values of the dimensions and the model results. We use the data from our test set - in cases where no data are available we use suitable examples. The analysis reveals that all the models react sensitively to all the dimensions both landmark and personal dimensions. We present the results for the models below.

- **PwSm**
  - **Landmark Dimensions** We found the highest sensitivity of the PwSm to the landmark dimensions. This sensitivity is not only dependent on the values of the attributes of the landmark dimensions of the investigated object but also on the values of the other objects at the decision point.
  - **Prior Spatial Knowledge** The sensitivity to *prior spatial knowledge* ranges between the sensitivity to the landmark dimensions and the sensitivity to *personal interests.* We identify a differentiation between $s_{PspK} \leq 3$ (meaning *prior spatial knowledge*) and $s_{PspK} = 7$ (meaning no *prior spatial knowledge*) for the PwSm.
  - **Personal Interests** The PwSm shows the lowest sensitivity to a variation in the inputs of the *personal interests.* However, the input values influenced the number of identified landmarks with the PwSm, meaning an interest rating ≤ 3 results in two, wheres a high interest rating > 3 results in three identified landmarks.

- **PwPm**
  - **Landmark Dimensions** The sensitivity of the PwPm to the landmark dimensions ranges between the sensitivity to $s_{pInt(gast)}$ and the other *personal interests.* It is dependent on the values of the attributes of the landmark dimensions of the other objects at the decision point, whether the model reacts sensitively to the landmark dimensions.
  - **Prior Spatial Knowledge** The sensitivity to the *prior spatial knowledge* ranges between the sensitivity to $s_{pInt(gast)}$ and the other *personal interests* but the SI is higher than the sensitivity for the landmark dimensions. It is the only model not confirming a differentiation between *prior spatial knowledge* and no *prior spatial knowledge.* Although it reacts sensitively to

the inputs of *prior spatial knowledge* it only makes a distinction between $s_{PspK} = 1$ and the other possible *prior spatial knowledge* ratings.

  - **Personal Interests** We observe the lowest sensitivity for the *personal interest* $s_{pInt(shop)}$, $s_{pInt(cult)}$, and $s_{pInt(hist)}$, whereas $s_{pInt(gast)}$ shows the highest observed sensitivity. The sensitivity analysis shows that a $s_{pInt} = 3$ (*medium*) results in different and, in addition, in a different number of identified landmarks than a $s_{pInt} = 4$ (*high*). However, $s_{pInt(shop)}$ is an exception for the PwPm. Although the model is sensitive to $s_{pInt(shop)}$, it makes no general distinction between *no* or *high personal interests* ratings.

- **PdFc**
  - **Landmark Dimensions** The sensitivity of the PdFc to s_{vis} and s_{sem} respectively is highly dependent on s_{PspK}. This is obvious because the model makes a clear distinction between *prior spatial knowledge* and no *prior spatial knowledge* (Figure 6.1). s_{str} does not appear in the model and, therefore, we cannot say anything about the sensitivity of s_{str}.
  - **Prior Spatial Knowledge** The PdFc is sensitive to the inputs of *prior spatial knowledge.* We identify a differentiation between $s_{PspK} \leq 3$ (meaning *prior spatial knowledge*) and $s_{PspK} \geq 6$ (meaning no *prior spatial knowledge*) for the PdFc.
  - **Personal Interests** The PdFc is sensitive to the *personal interests* because the model narrows down the identified landmarks according to the topics interesting for the traveller.

- **PdTm**
  - **Landmark Dimensions** The PdTm is sensitive to all landmark dimensions. It is dependent on the values of the attributes of the landmark dimensions of the investigated object whether a variation of the investigated dimension influences the output of the model. However, there is one exception: the PdTm identifies an object as a landmark in case the $s_{vis} > 69.758$ independent from the values of the other landmark dimensions.
  - **Prior Spatial Knowledge** The PdFc is sensitive to the inputs of *prior spatial knowledge.* It distinguishes between $s_{PspK} = 7$, meaning no familiarity at all, and the other ratings.

– **Personal Interests** The PdTm is sensitive to the pInt and splits either between a $s_{pInt} = 2$ (*low*) and $s_{pInt} = 3$ (*medium*) or between $s_{pInt} = 3$ (*medium*) and $s_{pInt} = 4$ (*high*) for most of the dimensions. Solely for $s_{pInt(gast)}$ the PdTm makes the distinction between $s_{pInt} = 5$ (*very high*) and the other ratings.

The survey participants did not select $s_{PspK} = 4$ and $s_{PspK} = 5$. In addition, the survey results show that on average only one survey participant chose $s_{PspK} = 6$ (Table 5.3). This suggests that these ratings do not influence the models. This becomes obvious for the PwSm and the PdTm since both models differentiate between a rating $s_{PspK} \leq 3$ (meaning *prior spatial knowledge*) and a rating $s_{PspK} = 7$ (meaning no *prior spatial knowledge*). We conclude that the PwSm and the PdTm follow a differentiation between participants familiar with the street intersection and all the others.

For most of the *personal interests* we identify a differentiation between $s_{pInt} = 2$ (rated *low*) or $s_{pInt} = 3$ (rated *medium*) and a $s_{pInt} = 4$ (rated *high*). The fact is that the *medium* rating is either assigned to the lower ratings or to the higher ratings. The strategy of choosing a midpoint can be explained by a phenomenon called *survey optimising* (Krosnick 1991). This behaviour occurs under cognitive load and when survey participants attempt to be fully diligent. As a consequence people sometimes try to avoid this effort but they want to answer responsibly (Krosnick 1991, Krosnick & Fabrigar 1997). Thus, the pInt rating *medium* might be either chosen by a participant who is actually interested as well as by a participant who is not.

7.2 Comparison of Model Results

We compare the results of the conventional models with the results of the personalised models to test our hypothesis. We count correctly identified landmarks and unidentified landmarks of the models and analyse the differences. For the comparison we use McNemar's test (Section 3.4). Table 7.22 shows the results - we discuss them in the following sections.

7.2.1 Conventional and Personalised Weighted Sum Model

The PwSm identifies the same number of landmarks as the CwSm. There are no landmarks changing from an unidentified to a correctly identified and also no landmarks changing vice versa (Table 7.22). Therefore, we are neither able to calculate

Table 7.22: Results of McNemar's test for the comparison of the conventional and the personalised models.

	PwSm/CwSm	PwPm/CwPm	PdFc/CdFc	PdTm/CdTm
Identified → Unidentified	0	0	11	33
Unidentified → Identified	0	20	2	36
Test Statistic	-	18.050	4.923	0.058
p-Value	-	<0.0001	0.027	0.810

a p-value nor a McNemar's test statistic. We conclude that there is no difference between these models and results can be considered as identical.

7.2.2 Conventional and Personalised Weighted Product Model

In order to find out whether there are differences between the PwPm and the CwPm we perform McNemar's test. We detect 20 discordant pairs all changing from an unidentified landmark with the PwPm to a correctly identified landmark with the CwPm (Table 7.22). The p-value is <0.0001 and McNemar's test statistic equals 18.050. This difference is considered to be extremely statistically significant. That means, there are no associations between the PwPm and the CwPm. We conclude that the PwPm identifies significantly less landmarks than the CwPm.

7.2.3 Conventional and Personalised Decision Flow Chart

Comparing the results of the CdFc and the PdFc, we find 13 discordant pairs (Table 7.22). There are 11 pairs where a landmark changes from an identified landmark with the PdFc to an unidentified with the CdFc and 2 pairs where a landmark changes vice versa. The p-value is calculated with a McNemar's test statistic of 4.923 and equals 0.027. This difference is considered to be statistically significant. This means, the PdFc identifies significantly more landmarks than the CdFc.

7.2.4 Conventional and Personalised Decision Tree Model

The CdTm identifies slightly more landmarks than the PdTm. To determine, whether this difference is significant we perform McNemar's test (Table 7.22). There are 69 discordant pairs, thereof 33 pairs where the PdTm identifies a landmark but the CdTm does not. The p-value equals 0.810 with a McNemar's test statistic of 0.058. By conventional criteria this difference is considered not to be statistically significant.

This means, that the CdTm does not identify significantly more landmarks than the personalised model.

7.2.5 Results of the Comparison Discussed

We compare the results of the personalised landmark identification models and the conventional models. It turns out that the CwSm identifies the same number of landmarks as the PwSm. The PwPm identifies - and that is extremely significant - less landmarks than the CwPm. We calculate the average of visual, semantic, and structural salience from the objects selected as landmarks by the survey participants. These average values are the basis of the initial weights for the PwSm and the PwPm. One issue with these weights is that some combinations of PspK and pInt ratings appear only once in the training set (compare Table A.3, Column *No*). We do not know how this has an affect on the results, but are aware that this might have an influence. One solution for this problem might be to consider only combinations of PspK and pInt ratings with an equal number of selections for the training of both models. The problem of this approach would be that we would not be able to identify many landmarks of the test set due to the missing initial weights of some combination of PspK and pInt ratings respectively. Therefore, for this initial investigation of personalised models, we do not consider the number of selections and use all the ratings available in both training and test set.

Another problem of the PwSm and the PwPm is that the PspK and the pInt are highly correlated. This mainly influences the search for optimal weights because the initial weights are dependent on both PspK and pInt. How to deal with that is still a field of research and methods have to be identified.

The PdFc is the only model that identifies - which is statistically significant - more landmarks than the corresponding conventional model. However, in absolute terms the PdFc shows the second-worst performance (Table 6.9). The CdFc even shows the worst performance of the conventional models (Table 5.9). The recall is less than one half of the recall of the CdTm. The poor CdFc result might be explained by the fact that we use a basic flow (Figure 4.5) because there is no other flow available. Maybe a modification of the CdFc similar to the training of the PdFc, as described in Section 6.1.2, would lead to better results. We built this basic flow to be able to compare the results of the PdFc to a model only considering landmark dimensions. We do not further evaluate or modify the flow of the CdFc as we did for the PdFc.

An issue of the PdFc is that the decision $s_{iLM} = s_{pInt} = 1$ (Figure 6.1) restricts the number of objects identified as landmarks. Thus, for a survey participant stating

$s_{pInt} = 0$ for all possible topics of interest this decision is not applicable and the objects passing the preceding process in the flow chart are all identified as landmarks. In most cases this results in more identified landmarks for a survey participant with no interests than for a survey participant stating an interest in a topic. That, in turn, has an impact on the number of landmarks identified by the PdFc because more identified landmarks mean a higher possibility that the one selected by the survey participant is amongst them.

The difference between the PdTm and the CdTm is not statistically significant and consequently, they identify approximately the same number of landmarks. The PdTm identifies a landmark for each decision point for each survey participant in this work. However, it might occur that the PdTm only identifies NALs for a decision point. One solution to solve this problem, is to stop the classification of an object not in the terminal leave but in another leave where it is still possible that the object becomes a landmark.

7.3 Conclusion from the Analyses and Comparison

In this chapter we performed a sensitivity analysis of our models to investigate the relationship between the input values of the dimensions (landmark and personal dimensions) and the model results. The analysis revealed that all the models react sensitively to all the dimensions. We compared the results of the personalised landmark identification models with the results of the conventional models. The PwSm, the PwPm, and the PdTm do not identify more landmarks than their corresponding conventional models. It turns out that the CwSm identifies the same number of landmarks as the PwSm. The PwPm identifies significantly less landmarks than the CwPm. The differences between the CdTm and the PdTm are not statistically significant. The only model that identified statistically more landmarks than the corresponding conventional model is the PdFc. However, in absolute terms this model shows the second-worst performance (compare Table 6.9). The recall of the CdTm is more than twice as high as the recall of the CdFc (Table 5.9). The reason for the poor result of the CdFc might be the basic flow (Figure 4.5) that we built as a CdFc is not existing. We conclude that the comparison of the results of PdFc and CdFc are not conclusive enough to confirm the hypothesis. Thus, although the personalised landmark identification models react sensitively to the personal dimensions, we have to reject the hypothesis that a personalised model considering prior spatial knowledge and personal interests identifies more landmarks selected by humans than a conventional model.

Chapter 8

Discussion of the Results

In Section 7.3 we concluded that we have to reject the hypothesis of this work. The most obvious interpretation for the rejection of our hypothesis is that the personal dimensions *prior spatial knowledge* and *personal interests* are not important for personalised landmark identification. However, there are a number of other reasons for this result. This chapter discusses five major points that might have influenced our result: missing other dimensions (Section 8.1), the methods to calculate salience values (Section 8.2), the models to calculate overall salience (Section 8.3), the dataset (Section 8.4), and the survey design (Section 8.5).

8.1 Further Dimensions

In this thesis we considered landmark dimensions as well as personal dimensions. However, there might be a number of other dimensions which might play a role for personalised landmark identification (Figure 8.1). These might be landmark dimensions influencing the underlying conventional models, personal dimensions influencing the personalised models, and other not yet identified dimensions.

8.1.1 Landmark Dimensions

In this thesis we considered the landmark dimensions visual, semantic, and structural dimensions. There might be a number of other important landmark dimensions influencing the underlying conventional models (Figure 8.1). During this thesis we came across two additional landmark dimensions: the *permanence* and the *descriptiveness* of an object.

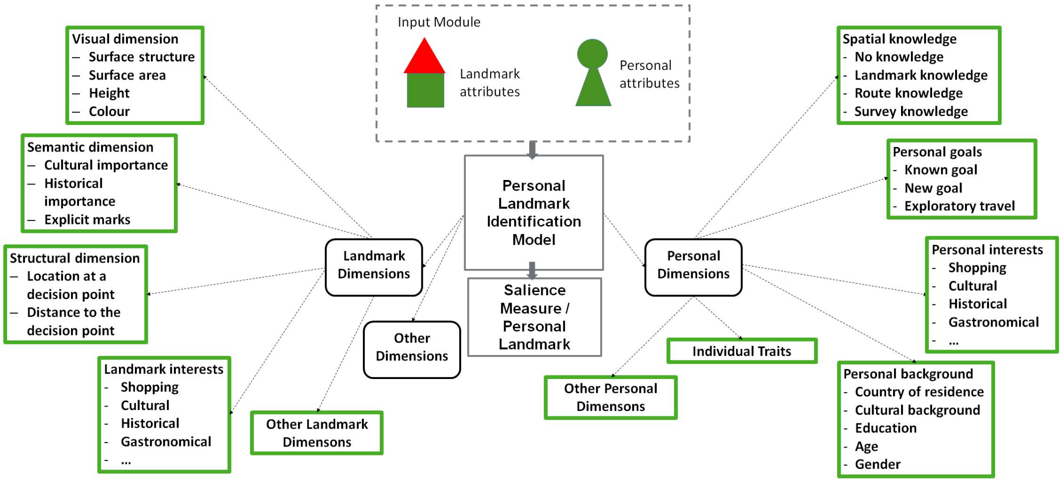


Figure 8.1: Dimensions of personalised landmark identification models.

Permanence

Burnett et al. (2001) suggest amongst other factors *permanence* as an important characteristic of landmarks. Studies investigate whether stable objects (fountains or monuments) are more informative than relatively unstable objects. Results show that there is a difference in degree how stable and unstable objects influence performance (Scrivner-Limbaugh 2015). Landmarks might change in various ways. They may be vacant objects (e.g. the building where the post office used to be) but they can also change their use (e.g. a clothes shop may become a restaurant or vice versa). Objects might change their size or appearance. For example we had the fountain at decision point 0 which changes its looks during winter (Figure 8.2a) and summer (Figure 8.2b). As during winter time it is just a box, survey participants did not really like this object, and we may assume that in summer time the result would be different. We cannot completely exclude that other objects did change during the survey, and might influence the results.

Descriptiveness

Another important dimension of a landmark is its descriptiveness. The *brevity* of a landmark description relates to the conciseness of the description needed to describe an object (Burnett et al. 2001). Richter & Duckham (2008) state that compact route directions together with landmarks are easier to understand. People may select a landmark because they know a brief and concise description for it, such as 'the coloured house' (DP 3, O_3, Figure A.4) compared to the 'grey small house with the dome roof' (DP 3, O_5, Figure A.4). Two messages with different content

(a) Winter. (b) Summer.

Figure 8.2: Non permanent object.

and length can communicate the same message, this concept is known in GIScience as *pragmatic semantics* (Frank 2003). Any object can be refereed to in a number of ways using different perspectives (Schober 1998). An example might be found in our introduction (Section 1): 'where once the mining director lived' and 'the house with the stucco façade' might refer to the same object, although these are completely different descriptions. There are findings that the names of POIs change with geographic distances (Hu & Janowicz 2018) suggesting that the description needed is dependent on the spatial knowledge of the wayfinder and may influence the selections of our survey participants.

8.1.2 Personal Dimensions

The identified personal dimensions are the basis for our calculations. We investigated *prior spatial knowledge*, *personal interests*, *personal goals*, *personal background*, and *individual traits*. Finally, we concentrated on *prior spatial knowledge* and *personal interests*. Here we give some ideas how the other not yet considered personal dimensions might have influenced our result. We are aware that there might be other not yet identified personal dimensions (Figure 8.1).

Personal Goals

Personal goals are already extensively discussed in Section 4.1.2. We divide three wayfinding tasks (Allen 1999): known goal, new goal, and exploratory travel. Personal goals might be included by adjusting the number and the distribution of landmarks along the route. The setting of our survey did not allow us to capture personal goals. All of the participants had the same goal: to complete the survey along the predefined route. Otherwise, we would need to change the survey settings to investigate the influence of personal goals (Section 8.5).

Personal background

There are a number of attributes discussed concerning the personal background in Section 4.1.2. We include questions on the background of the participants in our survey. In this thesis these data are only used in order to get an overview of the distribution of age, place of residence, and education of the participants. Nevertheless, the personal background of our survey participants might have influenced the results. Consider on the one hand participants living since their early childhood or birth in Augsburg and on the other hand participants that were in town before simply for shopping. Participants of both groups might state that they have been at the street intersection before and that they are familiar with the area. Are there differences in their behaviour of landmark selection?

There are hints that the age of the participants influences results. Participants older than 50 mentioned reasons often related to visibility for the selection of objects ('explicit mark is clearly visible', 'big letters and visible colour', or 'too small'). Younger participants mentioned visibility associated reasons but they were primarily connected to the walking direction ('small and out of sight' or 'visible from walking direction'). There might be other attributes of the personal background influencing our survey results.

Individual Traits

Unlike the other dimensions, individual traits can only be determined through specially designed psychological tests. For this reason, in this thesis individual traits are not further discussed although, they may affect results of the survey as it is for example known that emotions influence landmark selections (Section 2.3), which may be influenced by individual traits.

Other Personal Interests

In this thesis positive *personal interests* are treated. However, individual interests might be driven by the needs of a specific navigation task. Travellers might remember objects because a situation or personal need makes them pay attention even if it does not belong to the traveller's topic of interest. *Personal interests* are treated in a positive way in this thesis. However, interest may have neutral or even negative values. Landmarks part of negative topics are avoided during navigation, e.g. dirty or dangerous places in a city. Our interest is closely linked to emotions (Section 2.3). A place might have a high identifiability on a subjective basis because it is linked to negative experiences (e.g. this is the place where my car broke down).

However, our point is that there seems to be no influence on the survey results because of negative interest. Participants are asked to select one LM and one NAL for a personalised route direction. The participants comments showed that they associated the objects they did like and selected as a LM obviously with positive emotions ('like the architecture looks friendly', 'friends live here', or 'I like the beauties who are working here'), whereas objects they did not like were described with negative wordings ('low quality', 'boring house', or 'ugly building').

8.1.3 Other Dimensions

There might be other not yet identified dimensions besides landmark and personal dimensions (Figure 8.1). Two additional dimensions might be *environmental dimensions* and *context dimensions*.

Environmental Dimensions

The determination of landmarks for a specific route is known as *landmark integration* (Richter & Winter 2014). The focus is on environmental dimensions investigating the environment of the landmark as well as the route and the relationship to objects located nearby. This thesis investigates landmark identification models without such an environmental dimension. Nonetheless, there are hints that the environment of the objects as well as the whole route influenced landmark selections. Although we tried to avoid the influence of turning and walking directions by not letting the people know the itinerary route, participants always knew the approaching direction to the intersection. This knowledge seems to be reflected in the object selections because participants seem to choose landmarks simply because of their position. A number of participants state that they like the object because it is 'visible from walking direction', 'face to face when you come out of that street', or the 'center of view'. A reason mentioned repeatedly for dislike was that the participants 'have to look back' to see the object. These are environmentally dependent dimensions already extensively studied (Wang & Spelke 2000, Hollands et al. 2002, Röser 2015, Albrecht & von Stülpnagel 2018) appearing to influence the object selections in our survey.

Context Dimensions

There are several dimensions of context which might impact the results. One of them are weather conditions. We collected the data between October and February 2019. As it was wintertime, there were days when the objects along the route were covered with snow. Kattenbeck (2016) does not find severe differences between landmark salience ratings when objects are covered with snow and ratings without snow. Therefore, we assume that we may neglect the influence of snow. However, all objects were recognisable under snow and, if at all, their surface areas were covered only. We avoided the Christmas market because stands would hide the fountains and the statue and undertake no surveys during that time. All surveys were completed during day when there still was light.

8.2 Methods to Calculate Salience Values

We investigated and developed methods to calculate salience values for the landmark dimensions as well as for the personal dimensions *prior spatial knowledge* and *personal interests* and included them in our models. There might be other salience measures leading to more accurate model results.

8.2.1 Landmark Dimensions

This thesis assigned landmark salience values in percent to an object as soon as an attribute value was different or differed from the attribute values of the surrounding objects. In case all attribute values of a landmark dimension are salient, the object gets a 100% salience for this dimension. An object must fulfil specific conditions to be considered as salient (Table 4.1). We based our salience measures on threshold values from Raubal & Winter (2002) and Nuhn et al. (2012). They present - just as we do in this thesis - their salience measures without the empirical evidence that they lead to better results compared to other salience measures. Thus, the salience measures and the conditions which must be fulfilled for the attributes to be considered salient are based on many assumptions. These assumptions might be validated in the framework of a future empirical study.

8.2.2 Personal Dimensions

In this thesis we use the framework which Montello (1998) named the *dominant* framework (Siegel & White 1975) to measure *prior spatial knowledge* salience. Ishikawa & Montello (2006) identify the idea that landmark knowledge is a prerequisite for route knowledge, which again is mandatory for survey knowledge as a problem of this framework. As a solution they postulate different types of knowledge that are acquired simultaneously. This framework is referred to as the *continuous* framework. The survey participants did not select *prior spatial knowledge* saliences $s_{PspK} = 4$ and $s_{PspK} = 5$. In addition, the survey results show that on average only one survey participant chose $s_{PspK} = 6$ (Table 5.3). A continuous framework might be more useful to capture these stages of *prior spatial knowledge* and to differentiate more stages of no *prior spatial knowledge*.

Perhaps it would have sufficed to measure *prior spatial knowledge* salience only in two-stages: *prior spatial knowledge* and *no prior spatial knowledge* at the respective street intersections. This would be in line with Winter et al. (2005) who measured familiarity on a simple binary scale (but did not evaluate it further). The PdFc for example performs best when its flow contains only no *prior spatial knowledge* and *prior spatial knowledge* at a particular decision point instead of dividing in landmark, route, and survey knowledge. The PdTm and the PwSm provide hints and they make a distinction between participants who have never been at a particular street intersection before and the other participants. This might also be valid for the PwPm.

A weak point of the PdFc is that salience for *personal interests* is only measured in two ways: *interested* and *not interested*. It does not play a role whether the participant is interested in more than one topic of interest or just in one. Let us take two persons: the one is interested in gastronomy and shopping while the other is exclusively interested in gastronomy. Not withstanding their different interests, both situations are treated in the same way.

The salience of an object is dependent on its assignment to topics of interest. In order to be as objective as possible we used official databases. Furthermore, we exclusively used the four top ranked interest: shopping, cultural, historical, and gastronomy. However, we do not expect differences when considering all possible topics of interest. At our decision points there were ≤ 4 objects for the topics of interest not considered in this thesis (Table 5.1). The identification of effects would not be very specific and informative and results might be interpreted with difficulty. To investigate the models considering all the topics of interest, we would need an investigation area with more objects that are part of all possible different topics of interest.

8.3 Models to Calculate Overall Salience

All four tested models in this work are reasonable from a theoretical perspective. Other models or model settings might be useful as well and maybe would even deliver better results.

8.3.1 Underlying conventional Models

We decided to use a CwSm, a CwPm, a CdFc, and a CdTm as conventional models. During this thesis, it turned out that the conventional models are not as good in landmark identification as expected. As regards the models based on theory, their recalls for the training set are around 60% (Table 5.8), whereas their recalls for the test set merely reach around 40% (Table 5.9). We can conclude that none of the conventional models based on theory are good identifiers of landmarks.

As proposed by Raubal & Winter (2002) we used weights of one ($w_{vis} = w_{sem} = w_{str} = 1$) for the CwSm. To be able to compare the results of the CwPm with the results of the CwSm, we set its weights also to one. There are studies saying that different landmark dimensions have a different impact on successful landmark identification which outlines the importance of weighting each dimension relative to its significance (Kattenbeck 2016, Sadeghian & Kantardzic 2008). In future work it might be investigated whether the recalls of the CwSm and the CwPm can be improved by considering weights other than one.

The CdFc shows the worst performance of the conventional models. It reaches a recall of only 31.46% for the test set (Table 5.9). The poor CdFc result might be explained by the fact that we use a basic flow (Figure 4.5). Maybe a modification of the flow of the CdFc similar to the training of the PdFc, as described in Section 6.1.2, would lead to better results.

The CdTm is the only model that uses information from LMs as well as from NALs for training. It obtains an accuracy of 73.61% with the training set identifying LMs and NALs (Table 5.8, in brackets). The result differs, however, if we look only at landmarks resulting in a recall of only 56.75% (Table 5.8). The recall of the CdTm obtained with the test set is higher (Table 5.9) than the one obtained with the training set. This means the CdTm is better able to identify the landmarks in the test set than in the training set. The CdTm delivers the highest recall of all the conventional models and its result differs significantly from the results of the other models. Although the CdTm obtains the highest recall, there might be methods to improve its result.

8.3.2 Machine Learning Approaches

In addition to the models based on theory we investigated a machine learning model for landmark identification. We generally noticed a difference in test recall between this model (PdTm) and the models based on theory (PwSm, PwPm, PdFc). Maybe other machine learning approaches are suitable. In particular, if it comes to decision trees, one must admit the question: *Why do we not use a model based on random forest?* (Ho 1995, Breiman 2001). A random forest is essentially an algorithm that constructs a collection of decision trees. It randomly selects data entries and features in order to build multiple trees and then averages the results. After the creation of a number of trees, each tree chooses the class. The class that appears most often is the output of the random forest algorithm. Especially in the cases of a large number of data entries, the random forest achieves increased classification performance and delivers accurate and precise results (Ali et al. 2012). Random forest algorithms are opaque and act like a black box (Breiman 2001) and are not simple to interpret. The simplicity of explanations was taken as a prerequisite for our work and is one advantage of decision trees, encouraging us to apply them in this thesis.

8.3.3 Consideration of NALs

The machine learning models CdTm and PdTm are the only models that use NALs for the training. Currently NALs are not used for the training of the PwSm, the PwPm, and the PdFc. The machine learning model makes a clear classification of objects in LMs and NALs. For the models inspired by theory, the process is somehow different. They identify objects that are landmarks from a pool of objects at a decision point. However, this does not necessarily mean that the other objects are NALs. Nonetheless, these objects which are not identified as landmarks might be included in the training as well. It would be worth to investigate whether the consideration of the objects not identified as landmarks leads to a better performance of the models based on theory.

In this thesis we included only landmarks in the test set and omitted NALs. This ignores identified NALs and unidentified NALs (Section 3.2). The NALs are already considered in the training of the machine learning models and we calculate their *accuracy* (Formula 3.5). Identified and unidentified NALs could give us important additional information on the accuracy of the models obtained with the test set.

8.3.4 Weights for the Personalised Weighted Sum/Product Model

The PwSm and the PwPm have weights for the visual, the semantic, and the structural salience (w_{vis}, w_{sem}, and w_{str}). These weights are composed of initial relative weights and model parameters. We analysed the objects selected as landmarks by survey participants with different *personal interests* and *prior spatial knowledge* ratings. Then, we averaged their visual, semantic, and structural salience ($\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$, Table A.3) for each combination of ratings. We determined the initial relative weights from these averages which were then multiplied with the model parameters (Section 6.1.1).

Using the arithmetic average might have disadvantages because the average is sensitive to extreme values. Imagine, for example, the visual saliences of 25, 25, 25, 25, and 100. The sum of the five saliences is 200 and their average is 40. This does not necessarily tell us something about the traveller's preferences of visual salience. Therefore, the average might not be the best measure when there are extreme values in the dataset. For such a case a measure based on the median of the data or based on data distributions might be better alternatives. However, for this first approach on modelling personalised landmarks we based ourselves on the arithmetic average to calculate initial relative weights.

8.3.5 Global versus Local Rating

For the PwSm and the PwPm the sensitivity analysis shows that the sensitivity to the landmark dimensions is not only dependent on the values of the investigated object but also on the values of the other objects at the respective decision point. For example: the PwSm identifies an object (O1) with a salience measure of 1.25 as the most salient one for decision point one and an object (O2) with a salience measure of 2.5 as the most salient one for decision point two. Although their salience measures are completely different, with the salience measure of O1 only half as high as the salience measure of O2, they are both the most salient objects at their particular decision point. This could be a hint that a local salience measure would be more appropriate than the global one of the PwSm and the PwPm. Instead of taking absolute numbers, maybe we should work with rankings, making it obvious, that relative values are not to be taken as absolute ones. Such rankings might be useful measures for comparing the objects at a particular decision point to identify landmarks.

8.3.6 Overall Model

In this thesis we developed four personalised models for landmark identification. All of these models identified landmarks for different travellers with different *prior spatial knowledge* and different *personal interests*. But finally, the personalised models are either not able to identify more landmarks selected by survey participants than a conventional model or their performance is insufficient. We think it would be definitively worth to investigate whether an individual model for each survey participant identifies more landmark selections than one overall model. Survey participants might be influenced by individual intangible parameters resulting in individual landmark selections. This makes it difficult or even impossible to find one optimal individual personalised landmark identification model.

8.4 Dataset

Banko & Brill (2001) made a comparison among four different machine learning algorithms. They increased the training set size to millions and investigated the trained models. They concluded that 'the performance of learners can benefit significantly from much larger training sets' (Banko & Brill 2001, p. 32). Compared to their dataset, our dataset including 503 landmarks and the same number of NALs is relatively small. We discuss points which might be negatively influenced due to the small size of the dataset. Another point of discussion is that our dataset might contain fuzzy and uncertain data influencing our model results.

8.4.1 Dataset size

A point of discussion is the small size of the dataset. This might have an impact on our results and on their analyses. We discuss possible impacts in the following sections.

Impact on the Results of the Machine Learning Approach

The relatively small dataset size might have an influence on the results of the machine learning approach. There are studies confirming that the average accuracy of decision tree models that are built with CART increases with a bigger sample size (Sug 2009). However, increasing the training size does not necessarily lead to a better accuracy or recall of landmark identification. Training the decision tree models using additional training objects may lead to a high accuracy on the training set but to a less efficient one on the test set. This happens because the decision tree model might overfit the training set (Dietterich 1995) and to an extent that it is difficult for the tree to identify new unseen data.

We split the collected dataset of 503 LMs and the same number of NALs into a training and a test set. This leads to an even smaller amount of data available for the training of the model. However, in order to use the dataset as best as possible we applied 10-fold cross-validation. We divided the dataset into 10 folds each having 10% of the full set. We created the machine learning model on 9 training folds and calculated the model's recall on the remaining validation fold. We repeated this process 10 times using each fold once as a validation fold. This gave us 10 accuracies of the model. Putting all the results together, we used 100% of the training set to validate the model. This means for our training set for the P/CdTm, including 252 LMs and 252 NALs, that we end up with an average accuracy of our machine learning models based on 504 objects (even if objects are not used simultaneously for validating). However, in this thesis we used the collected dataset as best as possible. We propose that a next step would be to increase the size of the dataset and evaluate the outcomes of the models.

Impact on the Determination of Weights for the PwSm and the PwPm

The PwSm and the PwPm have weights for the visual, the semantic, and the structural salience (w_{vis}, w_{sem}, and w_{str}). These weights are composed of initial relative weights and model parameters. We analysed the objects selected as landmarks by survey participants with different *personal interests* and *prior spatial knowledge* ratings. We determined the average of visual, semantic, and structural salience ($\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$, Table A.3) for each combination of ratings. We determined the initial relative weights from these averages which were then multiplied with the model parameters (Section 6.1.1). Column *No* in Table A.3 shows that some combinations of *prior spatial knowledge* and *personal interests* ratings appear only once in the training set. We could have considered exclusively combinations of the ratings with approximately the same number of selections for the training and the determination of weights for both models. However, this would consequently lead to an even smaller size of the training set and also of the test set. Thus, it would not be possible to identify many landmarks of the test set due to the missing initial weights of some combinations of *prior spatial knowledge* and *personal interests* ratings respectively. One possibility to approach the problem of missing combinations of *prior spatial knowledge* and *personal interests* ratings could be to selectively collect the data from people with exactly these specific *personal interests* and that *prior spatial knowledge*. The main challenge of the approach, however, is to find such people. Thus, methods will have to be identified to overcome the problem of missing combinations of *prior spatial knowledge* and *personal interests* ratings.

Impact on the Sensitivity Analyses

For most cases we used examples from our test set to investigate the models' sensitivity to landmark and personal dimensions. In cases where no suitable data were available, we used appropriate examples. That concerns missing combinations of *prior spatial knowledge* and *personal interests* ratings forcing us to use other methods to be able to demonstrate the sensitivity of the models to the dimensions. As already stated above selectively collected data from people with this specific *personal interests* and *prior spatial knowledge* might be a solution for this problem.

8.4.2 Fuzzy and Uncertain Data

The analysis of complex relationships with mathematical models with a number of different dimensions is sometimes very vague and uncertain in many ways. Many features are interdependent features and cannot be evaluated by conventional measuring methods (Chen et al. 2011). In this thesis we do not consider fuzzy and uncertain data but are aware that they might influence our results. Gerla (2001) discusses the example of a red rose in the light of fuzzy logic. We can directly transfer this discussion to our topic and the claim that α is a red object. The colour of the object might not look exactly red. Then α is neither fully true nor fully false meaning it is neither zero nor one but for example 0.8 (Gerla 2001). This discussion might be extended to other attributes.

The personal dimensions are also affected by fuzziness. Due to perceptual differences between humans the information on *prior spatial knowledge* and *personal interests* of participants of the survey are affected by uncertainty. There are different facets of uncertainty involved (Gasós & Saffiotti 1999). These facets include mainly bad observations due to wrong self-assessment while filling out the questionnaire and the vagueness introduced by the use of our deterministic rating scales.

8.5 Survey Design

All data of personal dimensions were collected by a survey. We led people along an inner city route and asked them questions. One problem of the survey seems to be the influence of the walking direction. Survey participants did not know the route before, thereby the influence of the direction in which the route leads was excluded. However, the approach to the street intersection was always known. The reasons given by participants indicate that exactly that had an influence on their selections. The reasons were: 'directly visible from tram stops', 'object is not well visible', 'you have to look back'.

The group of objects selected for our survey might limit the participant's object selections. Although participants may like all the objects at the decision point, they have nevertheless to select an object they *do not like* as a NAL. There are hints in their comments that they had sometimes difficulties to decide. Their comments were: 'There's nothing. Can't relate.', 'all other objects are pretty nice.', or 'it's not that I dislike it completely'.

Another issue might be that the survey provides identical objects for both LMs and NALs. This results in objects that are selected as LM as well as for NALs. This again results in identical values for s_{vis}, s_{sem}, and s_{str} appearing for LMs as well as for NALs in the training set. As a consequence an object with a particular combination of s_{vis}, s_{sem}, and s_{str} classified in the training set as a NAL might appear in the test set as a LM. This might have an effect on the results of our models because it might result in a number of unidentified landmarks and, thus, keep the recall low.

It turned out that survey participants did not chose the $s_{PspK} = 4$ and $s_{PspK} = 5$. The results show that on average only one survey participant chose $s_{PspK} = 6$ (Table 5.3). In addition, the results of the sensitivity analyses suggest that these ratings do not influence the creation of the models. As a consequence we might neglect the ratings $s_{PspK} = 4$ to $s_{PspK} = 6$ for *prior spatial knowledge* and focus exclusively on $s_{PspK} = 7$ for no *prior spatial knowledge* at the street intersection. The other ratings concerning *prior spatial knowledge* at the street intersection ($s_{PspK} = 1$ to $s_{PspK} = 3$) may be maintained.

Chapter 9

Conclusions and Future Work

This chapter first summarises our research (Section 9.1). We present the results of this thesis and draw conclusions in Section 9.2. Based on our findings, we finally present our ideas for future work (Section 9.3). We end this thesis with some concluding remarks (Section 9.4).

9.1 Summary

In this work we investigated landmark identification models. We investigated personalised models and conventional models without personalisation. We hypothesised that a personalised model that incorporates *personal interests* and *prior spatial knowledge* identifies significantly more landmarks selected by humans than a conventional, non-personalised model. For testing the hypothesis we developed four personalised models for landmark identification and compared their outcomes to landmarks obtained from a survey as well as to the outcomes of conventional models.

We started this thesis by giving an overview on related work that deals with landmarks in human wayfinding, the modelling of landmarks for directions, and approaches used in the identification of personalised landmarks. This related research showed that a landmark is something individual for each traveller. Furthermore, the thesis showed that up-to-date there is only limited work on personalised landmark identification.

We continued our research by investigating mathematical models and analysis methods in general. We introduced three models based on theory (the wSm, the wPm, and a dFc) and one machine learning model (a dTm). We investigated the traditional machine learning approach for model training and testing and identified a method to 'train' the models based on theory. Furthermore, we identified methods for sensitivity analysis and the statistical evaluation of the results of the models.

We identified both, landmark as well as personal dimensions, as important factors to be considered in personalised landmark identification models. We built on existing landmark dimensions (visual, semantic, and structural) and added an additional *landmark interest* dimension to consider the topic of interest. We investigated *prior spatial knowledge*, *personal interests*, *personal goals*, *personal background*, and *individual traits* and identified attributes and attribute values for them. This thesis is a first approach towards the identification of personalised landmarks. A full elaboration of all five dimensions is beyond the scope of this work. Therefore, we focused on *prior spatial knowledge* and *personal interests*.

We investigated methods to calculate salience of the dimensions. We adapted existing salience measures from Raubal & Winter (2002) and Nuhn et al. (2012) for the landmark dimensions and introduced a new salience measure for the additional *landmark interest* dimension. In addition, we investigated new methods to calculate salience of the personal dimensions and developed methods to calculate salience of *personal interests* and *prior spatial knowledge*.

We investigated three models based on theory: a weighted sum model (wSm), a weighted product model (wPm), and a decision flow chart (dFc). In addition, we investigated a decision tree model (dTm) which is an approach in the field of machine learning. The models differ both in number of detected *LMs* and classification of *NALs*. wSm and wPm determine overall landmark salience measures. The result of the dFc is a number of landmarks. dTm provides a classification in LMs and NALs. wSm, wPm, and dFc do not make a clear statement on NALs. However, for our purpose a statement on the most personalised landmark was sufficient. We developed a conventional weighted sum model (CwSm), a conventional weighted product model (CwPm), a conventional decision flow chart (CdFc), and a conventional decision tree model (CdTm). In addition, we developed a personalised weighted sum model (PwSm), a personalised weighted product model (PwPm), a personalised decision flow chart (PdFc), and a personalised decision tree model (PdTm).

The next step was the implementation of the landmark identification models. We implemented all the models and methods using ESRI's ArcGIS 10.5.1 together with Python toolboxes using Python 2.7.13. In addition, we used several tools for data mining and data analysis.

A large part of this thesis was spent on data collection for landmark and personal dimensions along an innercity route in Augsburg. Landmark dimensions were extracted from official databases or acquired via a field survey focusing on the objects at the decision points of the route. In addition, landmark dimensions were collected for objects within the range of 100 meters as a basis for the salience calculations. The personal dimensions were collected with a survey using ESRIs Survey123. The survey contained questions about the *personal background*, the *personal interests*, and the *prior spatial knowledge* at the decision points of the route. Furthermore, we asked survey participants to select one object they like (*landmark* (LM)) and one object they *do not like* (*not a landmark* (NAL)) from a group of objects at each decision point for a personalised route direction. In total 51 participants completed the survey and gave information on their personal dimensions. The resulting dataset was used as input for our models.

The collected dataset was divided into a *training* and a *test set*. We established a 50:50 training testing set ratio as most suitable and we divided our dataset into two sets that do not overlap spatially. We applied the traditional machine learning approach for the training of the machine learning model. Inspired by this traditional approach, we 'trained' the models based on theory and identified weights of both, the PwSm and the PwSm, as well as an optimal flow of the PdFc.

After the models were fully created we used them to identify landmarks of our test set. We compared their identified landmarks with landmarks selected by the survey participants. To find the personalised landmark identification model which performs best, we compared their results with a subsequent McNemar's test.

We performed a sensitivity analysis to identify dimensions - landmark as well as personal dimensions - influencing the output of the personalised models. This local sensitivity analysis involving variation of only one dimension at a time analysed the effects on the models outputs.

For testing the hypothesis we focused on the comparison of the results of the personalised landmark identification models with the results of the conventional landmark identification models. We performed a McNemar's test to find out whether there are significant differences between the personalised landmark identification models and the conventional models.

9.2 Results and Conclusions

We investigated the results of the personalised landmark identification models and compared them to the results of the conventional models. Furthermore, we performed a sensitivity analysis. All models react sensitively to the landmark and the personal dimensions. However, PwSm, PwPm, and PdTm do not identify more landmarks than their corresponding conventional model. The PwPm even identifies significantly less landmarks than the CwPm. It turns out that the PdFc identifies statistically more landmarks than the CdFc. Nevertheless, in absolute terms the model shows the second-worst performance (compare Table 6.9). The recall of the CdFc is lower than the recalls of the other conventional models (compare Table 5.9). The recall of the CdFc is not even half as high as the one of the CdTm. The reason for the poor result of the CdFc might be the basic flow (Figure 4.5) that we built as a CdFc is not existing. For these reasons we conclude that the comparison of the results of PdFc and CdFc are not conclusive enough to confirm the hypothesis. We conclude that according to our results a personalised landmark identification model that incorporates *personal interests* and *prior spatial knowledge* does not identify significantly more landmarks selected by humans than a conventional, non-personalised model. Thus, we reject the hypothesis of this work. Our result confirms the findings of Gramann et al. (2017), Wunderlich & Gramann (2018) who also find that the additional effort of personalisation does not lead to improved results (Section 2.3). This shows that the data collection effort for obtaining information on *prior spatial knowledge* and *personal interests* for a pedestrian wayfinding application is unlikely to be justified.

9.3 Future Work

The investigation of possible reasons for rejecting the hypothesis in Section 9.2 revealed a number of open research questions. In this section we address the ones we think that they are worth further investigation.

How does the integration of further personal dimensions have an impact on our results? We want to introduce the personal dimensions excluded in this model in order to perhaps get a better landmark identification. Consequently, we need to research on how to model the other personal dimensions and how to calculate salience values. Especially the data of the *personal background* collected during this survey are worth to be investigated. These data might be analysed further to learn more about the relationship between *personal background* and the object selections of the individual participant. A further aspect is how to consider *personal goals*. We need to identify survey settings in order to get participants' wayfinding tasks with different wayfinding goals in mind. To be able to include *individual traits* in a personalised landmark identification model, we need to investigate them in depth to be able to identify possible modelling approaches. Psychological experts might be valuable assistants to provide knowledge and support in this area.

Does the inclusion of an environmental dimension in the models lead to better results? In our survey the environment of the objects as well as the route seem to influence the participants' object selections. There is already a lot of research regarding landmark integration (Section 2.2.2). We submit attributes and salience values for an environmental dimension in Nuhn & Timpf (2017*a*, 2018) for the analysis in a multidimensional model consisting of landmark, personal, and environmental dimensions. Attributes include *advance visibility* of an object (it always can be clearly seen from the route in all conditions (Burnett et al. 2001)), the traveller's *orientation* and *position* with respect to the object (Caduff & Timpf 2005*a*), and the *uniqueness* of an object, which cannot be mistaken for other objects and which is unique in its characteristics (Elias & Sester 2006). We make attempts to model these attributes in Nuhn & Timpf (2017*a*, 2018). However, such a multidimensional model still needs to be tested with real data, but research on the field should definitely be investigated further.

Does the application of other salience measures for the landmark dimensions have any affect on the results? In this thesis an object must fulfil specific conditions to be considered as salient. Salience is based on threshold values from Raubal & Winter (2002) and Nuhn et al. (2012). As soon as an object fulfils a salience condition because of a particular attribute value, it receives a percentage of a salience value for that particular attribute. Neither Raubal & Winter (2002) nor Nuhn et al. (2012), nor we provide an empirical evidence for any salience measures. Therefore, we do not know whether they are more suitable compared to other salience measures. However, we assume that the use of other salience measures might lead to different model results. Therefore, we recommend to conduct an empirical study applying different salience measures to identify those that yield the best model recall.

Does the application of the continuous instead of the dominant framework for prior spatial knowledge has any affect on the results? In this thesis we propose salience values for *prior spatial knowledge* as well as for *personal interests*. For *prior spatial knowledge* we base ourselves on what Montello (1998) named the 'dominant' framework (Siegel & White 1975). However, Montello (1998) proposes to use a continuous framework with different types of knowledge acquired simultaneously. The transfer of our salience values into a continuous framework is worthwhile, but needs a previous comprehensive analysis of spatial knowledge types.

Is it worth considering more topics of interest? In our model we consider the *personal interests*: *shopping*, *cultural*, *historical* and *gastronomy*. How do the results of our models differ as soon as we include the other topics of interest (Table 5.1)? To do so, we need another survey environment with more objects belonging to these topics of interest. In addition, we need guidelines for the classification of the objects. Experiments should be carried out whether e.g. people that are interested in shopping really prefer - as we assumed it especially for the PdFc - the shopping related things. However, not all persons want to go to the city for shopping but like the alternative to visit the internet. Such interest could be more differentiated and humans should be interviewed to find out what interest really means and how its consideration can contribute to a successful navigation.

Is it possible to improve the performance of conventional models? In this work we take data that are easy to obtain for landmarks dimensions. Unfortunately, the famous model of Raubal & Winter (2002) still needs to be tested in real life to have a basis of comparison for our results. Nothegger et al. (2004) provide a first approach but did not fully implement the whole model. They focus in their analysis only on one class of features in urban environments (façades) and investigate exclusively visual and semantic salience. As far as we know the overall model from Raubal & Winter (2002) uses slightly different attributes from ours. The results of the conventional models could be improved by adapting our visual, semantic, and structural attributes. Furthermore, an additional effort might be spent on the collection and prepossessing of data as basis for the calculation of the salience values of objects. In addition, there are studies outlining the importance weighting the dimensions relative to their significance in the conventional models (Kattenbeck 2016, Sadeghian & Kantardzic 2008, Kattenbeck et al. 2018). Further tests should be carried out to examine whether the introduction of weights leads to better results of the CwSm and the CwPm.

At the moment the CdFc consists of three processes based on the visual, the semantic, and the structural salience of an object. We could for example include processes based on attributes of the dimensions to find out whether this improves the recall of the CdFc.

One possible method to improve the performance of the machine learning model might be feature selection (Stein et al. 2005, Sugumaran et al. 2007, Rao et al. 2019). During feature selection a subset of relevant dimensions for the use in constructing the model is selected. Generally not all dimensions are relevant (Stein et al. 2005) and some of them might have a higher impact than others (compare results of the sensitivity analyses, Section 7.1). However, the selection of good dimensions requires detailed domain knowledge (Sugumaran et al. 2007). It might be worth to investigate whether the selection of only a few dimensions to train our machine learning model might lead to a better recall then.

Does the inclusion of NALs in all the models lead to better results? Currently, the machine learning models CdTm and PdTm are the only models that use NALs for the training. The other models based on theory neglect this information. We think it would be definitively worth to investigate whether the inclusion of NALs in the training set helps us to improve the models. Objects which are identified as NALs with the models could be compared with the NALs selected by survey participants. In addition to the *recall* of the model, we could then calculate its *accuracy* (Formula 3.5) giving us additional information on the model's performance. Additional performance measures based on NALs such as the *precision* would be possible (Buckland & Gey 1994). The *precision* is the ratio of the *Identified LMs* and the sum of *Identified LMs* and *Unidentified NALs* (Table 3.2). It shows how much the model correctly identifies a landmark out of all the objects which the model identifies as a landmark (Buckland & Gey 1994). These measures might help us to improve the models both machine learning and models based on theory.

Is one individual model for each traveller more suitable than an overall model? In this thesis we chose to build one overall model incorporating the survey results from all participants. Another approach would be to build one individual model for each participant or for groups of participants with similar combinations of *prior spatial knowledge* and *personal interests* ratings. Our approach is more generalisable and avoids overfitting to one participant. However, landmark selections might vary considerably between humans, therefore, one model for each traveller might result in a higher recall. In future work one model might be created for each participant and the performance of these individual models might be compared. It might be assessed how the individual models may fit for a traveller or a group of travellers and their performance might be compared to our overall model.

Is it more suitable to apply a larger dataset? We identified the small size of the collected dataset as a weak point of this thesis. It might have a strong impact on the results of the machine learning approach, on the weights of the PwSm and the PwPm, and on the sensitivity analyses. In addition, there might be various other impacts on the model results. Increasing the sample size would be worth a further investigation. There are several methods to increase it: collecting additional data along our route, identifying other suitable datasets, or artificially increasing the number of training samples. In his experiment Kattenbeck (2016) collected demographic data such as the *background* of a traveller and the *knowledge about a place* (Section 2.3). Kattenbeck et al. (2018) transferred his model from Regensburg to Augsburg. Results showed that the relationships between the subdimensions of salience do not differ significantly in the other city. Therefore, we may assume that the transfer of collected data from Regensburg to Augsburg is possible. For this reason, it would be worth to investigate whether Kattenbeck's (2016) dataset of Regensburg as well as the dataset of Kattenbeck et al. (2018) from Augsburg might be used to extend our collected dataset. The *knowledge about a place* might be mapped to our *prior spatial knowledge* and maybe this leads to additional combinations of *prior spatial knowledge* and *personal interests.*

Is it more suitable to apply a fuzzy method to model the landmark and personal dimensions instead of a deterministic one? The fuzzy set theory is introduced by Zadeh (1965). This idea is used in many analysis models to solve fuzzy problems (Mardani et al. 2015). There are approaches interpreting decision trees by using fuzzy logic (Bhalchandra et al. 2015, Mendonça et al. 2007, Cintra et al. 2013). There exists research on flow charts and fuzzy sets (Tanaka & Mizumoto 1975, Ostasiewicz 1982) associating decisions with a fuzzy relation and a fuzzy assignment. Input, outputs, and decisions might represent fuzzy sets. Gasós & Saffiotti (1999) discuss techniques to represent and use uncertain spatial knowledge in the field of autonomous robotics which might be worth to be evaluated and transferred to our use case.

What kind of results deliver other survey settings? The resulting data of our survey might be a useful resource for further studies on both modelling personalised landmarks as well as understanding preferences of travellers. However, there are a number of possible ways to change the survey settings and to investigate the results. What are the results when we first identify the landmarks with our models and then present it to survey participants? They could rate how they like it, for example on some rating scale. From the results of such a survey performance measures for the models might be derived. What happens when survey participants choose intersections part of different stages of their spatial knowledge and hosting objects interesting for them? Participants could be told to select for each stage of *prior spatial knowledge* a street intersection. In addition, participants might be directly acquired for the survey according to their level of interest in a topic. Additionally, a higher number of survey participants may supply more data. Such an approach might be especially useful to overcome the problem of the missing combinations of *prior spatial knowledge* and *personal interests* ratings. What objects do participants select when they are completely free? Participants may choose LMs and NALs completely free without doing preselections. This might overcome the problem of having to select an object although there is none that they either like or dislike.

9.4 Final Remarks

This thesis is a first approach on modelling personalised landmarks. In Chapter 1 we state that the highest cost for the provision of personalised landmarks is the personal data collection. Furthermore, we outlined that we need to be sure that the data collection effort is justified in relation to the advantages that can be achieved through the provision of personalised landmarks. We found out that *prior spatial knowledge* and *personal interests* play a role for the identification of personalised landmarks and that the personalised models react sensitively to their input values. We showed that a personalised landmark identification model that incorporates *prior spatial knowledge* and *personal interests* does not identify more landmarks selected by survey participants than a conventional, non-personalised model. The most obvious interpretation for this finding is that these personal dimensions are not important for landmark identification. We discussed a number of other reasons for this result and revealed open research questions. However, we currently have to conclude that the data collection effort for obtaining information on *prior spatial knowledge* and *personal interests* for an applied system might not be justifiable. In case future research confirms our findings it is most likely sufficient to focus on existing conventional, non-personalised models and to concentrate on their use in applied pedestrian wayfinding applications.

Bibliography

Albrecht, R. & von Stülpnagel, R. (2018), Memory for salient landmarks: Empirical findings and a cognitive model, *in* S. H. Creem-Regehr, J. Schöning & A. Klippel, eds, 'Spatial Cognition XI - 11th International Conference, Spatial Cognition 2018', Vol. 11034 of *Lecture Notes in Computer Science*, Springer, pp. 311–323.

Ali, J., Khan, R., Ahmad, N. & Maqsood, I. (2012), 'Random forests and decision trees', *International Journal of Computer Science Issues* **9**(5), 272.

Allen, G. L. (1997), From knowledge to words to wayfinding: Issues in the production and comprehension of route directions, *in* S. C. Hirtle & A. U. Frank, eds, 'Spatial Information Theory. A Theoretical Basis for GIS. COSIT 1997', Vol. 1329 of *Lecture Notes in Computer Science*, Springer, pp. 363–372.

Allen, G. L. (1999), Spatial abilities, cognitive maps, and wayfinding, *in* R. G. Golledge, ed., 'Wayfinding behavior: Cognitive mapping and other spatial processes', Vol. 4680, Johns Hopkins University Press Baltimore, pp. 46–80.

Allen, G. L. (2000), 'Principles and practices for communicating route knowledge', *Applied Cognitive Psychology* **14**(4), 333–359.

Allen, G. L. & Kirasic, K. C. (2003), Visual attention during route learning: A look at selection and engagement, *in* W. Kuhn, M. F. Worboys & S. Timpf, eds, 'Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003', Vol. 2825 of *Lecture Notes in Computer Science*, Springer, pp. 390–398.

Amari, S.-I., Murata, N., Muller, K.-R., Finke, M. & Yang, H. H. (1997), 'Asymptotic statistical theory of overtraining and cross-validation', *IEEE Transactions on Neural Networks* **8**(5), 985–996.

Appleyard, D. (1969), 'Why buildings are known: A predictive tool for architects and planners', *Environment and Behavior* **1**(2), 131–156.

Apté, C. & Weiss, S. (1997), 'Data mining with decision trees and decision rules', *Future Generation Computer Systems* **13**(2), 197 – 210.

Arthur, P. & Passini, R. (1992), *Wayfinding: people, signs, and architecture*, McGraw-Hill Book Company.

ASME (1947), *ASME standard operation and flow process charts*, The American society of mechanical engineers.

Azar, F. S. (2000), *Multiattribute decision-making: use of three scoring methods to compare the performance of imaging techniques for breast cancer detection*, Technical Report, Department of Computer & Information Science, University of Pennsylvania.

Bahn, V. & McGill, B. J. (2013), 'Testing the predictive performance of distribution models', *Oikos* **122**(3), 321–331.

Baird, B. F. (1989), *Managerial decisions under uncertainty: An introduction to the analysis of decision making*, John Wiley & Sons, Inc.

Balaban, C. Z., Karimpur, H., Röser, F. & Hamburger, K. (2017), 'Turn left where you felt unhappy: how affect influences landmark-based wayfinding', *Cognitive Processing* **18**(2), 135–144.

Balaban, C. Z., Röser, F. & Hamburger, K. (2014), The effect of emotions and emotionally laden landmarks on wayfinding, *in* 'Proceedings of the Annual Meeting of the Cognitive Science Society', Cognitive Science Society, pp. 1880–1885.

Banerjee, S., Frey, H.-P., Molholm, S. & Foxe, J. J. (2015), 'Interests shape how adolescents pay attention: the interaction of motivation and top-down attentional processes in biasing sensory activations to anticipated events', *European Journal of Neuroscience* **41**(6), 818–834.

Banko, M. & Brill, E. (2001), Scaling to very very large corpora for natural language disambiguation, *in* B. L. Webber, ed., 'Proceedings of the 39th annual meeting on association for computational linguistics', Association for Computational Linguistics, pp. 26–33.

Bayerisches Landesamt für Denkmalpflege (2018), 'Bayerischer Denkmalatlas'. Accessed November 2018.
URL: *geoportal.bayern.de/bayernatlas-klassik/*

Berry, D. C. & de Rosis, F. (1991), Designing an adaptive interface for epiaim, *in* S. Mario, H. Arie, F. Marius & T. Jan, eds, 'AIME 91', Springer, pp. 306–316.

Bhalchandra, P., Khamitkar, S., Deshmukh, N., Lokhande, S. & Mekewad, S. (2015), Characterization of fuzzy tree searches: A perspective note, *in* D. Mandal, R. Kar, S. Das & B. K. Panigrahi, eds, 'Intelligent Computing and Applications', Vol. 343, Springer, pp. 1–10.

Borra, S. & Di Ciaccio, A. (2010), 'Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods', *Computational Statistics & Data Analysis* **54**(12), 2976–2989.

Boulos, M. N. K. (2005), 'Web GIS in practice III: creating a simple interactive map of England's Strategic Health Authorities using Google Maps API, Google Earth KML, and MSN Virtual Earth Map Control', *International Journal of Health Geographics* **4**(22), 8.

Bradford, J. P., Kunz, C., Kohavi, R., Brunk, C. & Brodley, C. E. (1998), Pruning decision trees with misclassification costs, *in* C. Nedellec & C. Rouveirol, eds, 'European Conference on Machine Learning', Springer, pp. 131–136.

Bramer, M. (2007), *Principles of data mining*, Springer.

Breiman, L. (2001), 'Random forests', *Machine Learning* **45**(1), 5–32.

Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. (1984), *Classification and Regression Trees*, Taylor & Francis.

Bridgman, P. (1922), *Dimensionless Analysis*, New Haven CT: Yale University Press.

Brodley, C. E. & Utgoff, P. E. (1995), 'Multivariate decision trees', *Machine Learning* **19**(1), 45–77.

Brownlee, J. (2019), 'How to calculate Mcnemar's test to compare two machine learning classifiers'. Accessed January 2019.
URL: *https://machinelearningmastery.com/mcnemars-test-for-machine-learning/*

Brunk, C. A. & Pazzani, M. J. (1991), An investigation of noise-tolerant relational concept learning algorithms, *in* L. A. Birnbaum & G. C. Collins, eds, 'Machine Learning Proceedings 1991', Elsevier, pp. 389–393.

Brusilovsky, P. & Millán, E. (2007), User models for adaptive hypermedia and adaptive educational systems, *in* B. Peter, K. Alfred & N. Wolfgang, eds, 'The adaptive web', Springer, pp. 3–53.

Buckland, M. & Gey, F. (1994), 'The relationship between recall and precision', *Journal of the American society for information science* **45**(1), 12–19.

Budiharjo, A. P. W. & Abulwafa, M. (2017), 'Comparison of weighted sum model and multi attribute decision making weighted product methods in selecting the best elementary school in indonesia', *International Journal of Software Engineering and Its Applications* **11**(4), 69–90.

Burnett, G., Smith, D. & May, A. (2001), Supporting the navigation task: characteristics of 'good' landmarks, *in* M. A. Hanson, ed., 'Contemporary Ergonomics 2001', Taylor & Francis, pp. 441–446.

Caduff, D. & Timpf, S. (2005*a*), The landmark spider: Representing landmark knowledge for wayfinding tasks, *in* T. Barkowsky, C. Freksa, M. Hegarty & R. Lowe, eds, 'AAAI spring symposium: Reasoning with mental and external diagrams: Computational modeling and spatial assistance', AAAI Press, pp. 30–35.

Caduff, D. & Timpf, S. (2005*b*), The landmark spider: Weaving the landmark web, *in* '5th Swiss Transport Research Conference', Institute for Transport Planning and Systems, ETH Zurich, p. 14.

Caduff, D. & Timpf, S. (2008), 'On the assessment of landmark salience for human navigation', *Cognitive Processing* **9**(4), 249–267.

Cambridge Coding Academy (2019), 'Scanning hyperspace: how to tune machine learning models'. Accessed July 2019.
URL: *https://cambridgecoding.wordpress.com/2016/04/03/scanning-hyperspace-how-to-tune-machine-learning-models/*

Cambridge Dictionary (2019), 'Definition of personalization'. Accessed July 2019.
URL: *https://dictionary.cambridge.org/dictionary/english/personalization*

Chandrasekara, P., Mahaulpatha, T., Thathsara, D., Koswatta, I. & Fernando, N. (2016), 'Landmarks based route planning and linear path generation for mobile navigation applications', *Spatial Information Research* **24**(3), 245–255.

Chen, J., Wu, Z., Gao, H., Zhang, C., Cao, X. & Li, D. (2013), Recommending interesting landmarks based on geo-tags from photo sharing sites, *in* X. Lin, Y. Manolopoulos, D. Srivastava & G. Huang, eds, 'Web Information Systems Engineering – WISE 2013', Springer, pp. 151–159.

Chen, V. Y., Lien, H.-P., Liu, C.-H., Liou, J. J., Tzeng, G.-H. & Yang, L.-S. (2011), 'Fuzzy MCDM approach for selecting the best environment-watershed plan', *Applied Soft Computing* **11**(1), 265 – 275.

Chicco, D. (2017), 'Ten quick tips for machine learning in computational biology', *BioData Mining* **10**(1), 17.

Choi, J., McKillop, E., Ward, M. & L'Hirondelle, N. (2006), 'Sex-specific relationships between route-learning strategies and abilities in a large-scale environment', *Environment and Behavior* **38**(6), 791–801.

Cintra, M. E., Monard, M. C. & Camargo, H. A. (2013), 'A fuzzy decision tree algorithm based on C4. 5', *Mathware & Soft Computing Magazine* **20**(1), 56–62.

Coluccia, E. & Louse, G. (2004), 'Gender differences in spatial orientation: A review', *Journal of Environmental Psychology* **24**(3), 329–340.

Couclelis, H., Golledge, R. G., Gale, N. & Tobler, W. R. (1987), 'Exploring the anchor-point hypothesis of spatial cognition', *Journal of Environmental Psychology* **7**(2), 99 – 122.

Credé, S., Fabrikant, S. I., Thrash, T. & Hölscher, C. (2017), Do Skyscrapers Facilitate Spatial Learning Under Stress? On the Cognitive Processing of Global Landmarks, *in* F. Paolo, B. Andrea & C. Eliseo, eds, 'Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory', Springer, pp. 27–29.

Dabbs Jr, J. M., Chang, E.-L., Strong, R. A. & Milun, R. (1998), 'Spatial ability, navigation strategy, and geographic knowledge among men and women', *Evolution and human behavior* **19**(2), 89–98.

Daniel, M.-P. & Denis, M. (1998), 'Spatial descriptions as navigational aids: A cognitive analysis of route directions', *Kognitionswissenschaft* **7**(1), 45–52.

Denis, M., Pazzaglia, F., Cornoldi, C. & Bertolo, L. (1999), 'Spatial discourse and navigation: An analysis of route directions in the city of Venice', *Applied Cognitive Psychology* **13**(2), 145–174.

Dietterich, T. (1995), 'Overfitting and undercomputing in machine learning', *ACM Computing Surveys* **27**(3), 326–327.

Dietterich, T. G. (1998), 'Approximate statistical tests for comparing supervised classification learning algorithms', *Neural Computation* **10**(7), 1895–1923.

Dijkstra, E. W. (1959), 'A note on two problems in connexion with graphs', *Numerische Mathematik* **1**(1), 269–271.

Downs, R. M. & Stea, D. (1974), *Image and environment: Cognitive mapping and spatial behavior*, Aldine Transaction.

Dräger, M. & Koller, A. (2012), Generation of landmark-based navigation instructions from open-source data, *in* W. Daelemans, ed., 'Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics', Association for Computational Linguistics, pp. 757–766.

Duckham, M., Winter, S. & Robinson, M. (2010), 'Including landmarks in routing instructions', *Journal of Location Based Services* **4**(1), 28–52.

Edwards, A. L. (1948), 'Note on the "correction for continuity" in testing the significance of the difference between correlated proportions', *Psychometrika* **13**(3), 185–187.

Elias, B. (2003*a*), Determination of landmarks and reliability criteria for landmarks, *in* '5th Workshop on Progress in Automated Map Generalization', ICA Commission on Map Generalization.

Elias, B. (2003*b*), Extracting landmarks with data mining methods, *in* W. Kuhn, M. F. Worboys & S. Timpf, eds, 'Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003', Vol. 2825 of *Lecture Notes in Computer Science*, Springer, pp. 375–389.

Elias, B. (2006), Extraktion von Landmarken für die Navigation, PhD thesis, University of Hanover, Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Universität Hannover.

Elias, B. & Sester, M. (2006), Incorporating landmarks with quality measures in routing procedures, *in* M. Raubal, H. J. Miller, A. Frank & M. F. Goodchild, eds, 'Geographic Information Science. GIScience 2006', Vol. 4197 of *Lecture Notes in Computer Science*, Springer, pp. 65–80.

ESRI (2018), 'Python and arcpy'. Accessed November 2018.
URL: *https://desktop.arcgis.com/en/arcmap/latest/analyze/python/importing-arcpy.htm*

Everitt, B. S. (1992), *The analysis of contingency tables*, Chapman and Hall/CRC.

Fiacco, A. V. (1983), *Introduction to sensitivity and stability analysis in nonlinear programming*, Elsevier.

Fink, B. (1991), Interest development as structural change in person-object relationships, *in* L. Oppenheimer & J. Valsiner, eds, 'The Origins of Action', Springer, pp. 175–204.

Fishburn, P. (1967), 'Additive utilities with incomplete product set: applications to priorities and sharings google scholar', *Operations Research* **15**(3), 537–542.

Fontaine, S. & Denis, M. (1999), The production of route instructions in underground and urban environments, *in* C. Freksa & D. M. Mark, eds, 'Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. COSIT 1999', Vol. 1661 of *Lecture Notes in Computer Science*, Springer, pp. 83 –94.

Frank, A. U. (2003), Pragmatic information content - how to measure the information in a route description, *in* M. Duckham, M. F. Goodchild & M. F. Worboys, eds, 'Foundations of geographic information science', Taylor & Francis, pp. 47–70.

Fryman, M. A. (2002), *Quality and process improvement*, Delmar, Thomson Learning.

Fürnkranz, J. (1994*a*), A comparison of pruning methods for relational concept learning, *in* U. M. Fayyad & R. Uthurusamy, eds, 'AAAI.94 Workshop on Knowledge Discovery in Databases', pp. 371–382.

Fürnkranz, J. (1994*b*), Fossil: A robust relational learner, *in* F. Bergadano & L. De Raedt, eds, 'Machine Learning: ECML-94', Springer, pp. 122–137.

Fürnkranz, J. & Widmer, G. (1994), Incremental reduced error pruning, *in* W. W. Cohen, ed., 'Machine Learning Proceedings 1994', Elsevier, pp. 70–77.

Galea, L. A. & Kimura, D. (1993), 'Sex differences in route-learning', *Personality and individual differences* **14**(1), 53–65.

Gasós, J. & Saffiotti, A. (1999), 'Using fuzzy sets to represent uncertain spatial knowledge in autonomous robots', *Spatial Cognition and Computation* **1**(3), 205–226.

Geisser, S. (1975), 'The predictive sample reuse method with applications', *Journal of the American Statistical Association* **70**(350), 320–328.

Gelfand, S. B., Ravishankar, C. & Delp, E. I. (1989), An iterative growing and pruning algorithm for classification tree design, *in* 'Conference Proceedings., IEEE International Conference on Systems, Man and Cybernetics', IEEE, pp. 818–823.

Gerla, G. (2001), Probabilistic fuzzy logics, *in* R. Wójcicki, D. Mundici, K. Segerberg, A. Urquhart, H. Wansing & J. Malinowski, eds, 'Fuzzy Logic: Mathematical Tools for Approximate Reasoning', Springer, pp. 171–198.

Gilbreth, F. B. & Gilbreth, L. M. (1921), *Process charts*, American Society of Mechanical Engineers.

Golledge, R. (1991), Cognition of physical and built environments, *in* T. Gärling & G. W. Evans, eds, 'Environment, cognition, and action: An integrated approach', Oxford University Press on Demand, pp. 35–62.

Golledge, R. (1999), *Wayfinding behaviour*, The Johns Hopkins University Press Baltimore.

Golledge, R. G., Jacobson, R. D., Kitchin, R. & Blades, M. (2000), 'Cognitive maps, spatial abilities, and human wayfinding', *Geographical Review of Japan* **73**(2), 93–104.

Goodchild, M. F. (2007), 'Citizens as sensors: the world of volunteered geography', *GeoJournal* **69**(4), 211–221.

Goodman, J., Brewster, S. A. & Gray, P. (2005), 'How can we best use landmarks to support older people in navigation?', *Behaviour & Information Technology* **24**(1), 3–20.

Goren-Bar, D., Graziola, I., Pianesi, F. & Zancanaro, M. (2006), 'The influence of personality factors on visitor attitudes towards adaptivity dimensions for mobile museum guides', *User Modeling and User-Adapted Interaction* **16**(1), 31–62.

Götze, J. & Boye, J. (2013), Deriving salience models from human route directions, *in* Kelleher, R. John, Robert & S. Dobnik, eds, 'Proceedings of the IWCS 2013 Workshop on Computational Models of Spatial Language Interpretation and Generation', Association for Computational Linguistics, pp. 7–12.

Götze, J. & Boye, J. (2016), 'Learning landmark salience models from users' route instructions', *Journal of Location Based Services* **10**(1), 47–63.

Graham, B. B. (2004), *Detail process charting: speaking the language of process*, John Wiley & Sons, Inc.

Gramann, K., Hoepner, P. & Karrer-Gauss, K. (2017), 'Modified navigation instructions for spatial navigation assistance systems lead to incidental spatial learning', *Frontiers in Psychology* **8**, 193.

Gruber, M. J., Gelman, B. D. & Ranganath, C. (2014), 'States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit', *Neuron* **84**(2), 486–496.

Gupta, B., Rawat, A., Jain, A., Arora, A. & Dhami, N. (2017), 'Analysis of various decision tree algorithms for classification in data mining', *International Journal of Computer Applications* **8**, 15–19.

Guyon, I. (1997), A scaling law for the validation-set training-set size ratio, *in* 'AT & T Bell Laboratories', Citeseer, p. 11.

Hamburger, K. & Röser, F. (2014), 'The role of landmark modality and familiarity in human wayfinding', *Swiss Journal of Psychology* **73**(4), 205–213.

Hamby, D. (1994), 'A review of techniques for parameter sensitivity analysis of environmental models', *Environmental monitoring and assessment* **32**(2), 135–154.

Hamby, D. (1995), 'A comparison of sensitivity analysis techniques', *Health physics* **68**(2), 195–204.

Han, J. & Lee, H. (2015), 'Adaptive landmark recommendations for travel planning: Personalizing and clustering landmarks using geo-tagged social media', *Pervasive and Mobile Computing* **18**, 4 – 17.

Hart, P. E., Nilsson, N. J. & Raphael, B. (1968), 'A formal basis for the heuristic determination of minimum cost paths', *IEEE Transactions on Systems Science and Cybernetics* **4**(2), 100–107.

Hay, A. (1988), 'The derivation of global estimates from a confusion matrix', *International Journal of Remote Sensing* **9**(8), 1395–1398.

Heath, D., Kasif, S. & Salzberg, S. (1993), Induction of oblique decision trees, *in* R. Bajcsy, ed., 'Proceedings of the 13th International Joint Conference on Artificial Intelligence', Morgan Kaufmann Publisher, pp. 1002–1007.

Herrmann, T., Schweizer, K., Janzen, G. & Katz, S. (1998), 'Routen- und Überblickswissen - konzeptuelle Überlegungen', *Kognitionswissenschaft* **7**(4), 145–165.

Hidi, S. & Renninger, K. A. (2006), 'The four-phase model of interest development', *Educational psychologist* **41**(2), 111–127.

Ho, T. K. (1995), Random decision forest, *in* 'Proceedings of the 3rd International Conference on Document Analysis and Recognition', IEEE, pp. 14–18.

Hoffman, F. & Gardner, R. (1983), Evaluation of uncertainties in environmental radiological assessment models, *in* J. E. Till & R. H. Meyer, eds, 'Radiological Assessments: A Textbook on Environmental Dose Assessment', Report No. NUREG/CR-3332. US Nuclear Regulatory Commission, Washington, DC, pp. 11:1–11:55.

Hollands, M. A., Patla, A. E. & Vickers, J. N. (2002), '"look where you're going!": gaze behaviour associated with maintaining and changing the direction of locomotion', *Experimental brain research* **143**(2), 221–230.

Homma, T. & Saltelli, A. (1996), 'Importance measures in global sensitivity analysis of nonlinear models', *Reliability Engineering & System Safety* **52**(1), 1–17.

Howe, J. (2006), 'The rise of crowdsourcing', *Wired Magazine* **14**(6), 1–5.

Hsu, C.-W., Chang, C.-C., Lin, C.-J. et al. (2016), 'A practical guide to support vector classification'. Accessed August 2018.
URL: *https://www.csie.ntu.edu.tw/ cjlin/papers/guide/guide.pdf*

Hu, Y. & Janowicz, K. (2018), An empirical study on the names of points of interest and their changes with geographic distance, *in* S. Winter, M. Sester & A. L. Griffin, eds, '10th International Conference on Geographic Information Science (GIScience 2018)', Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik.

Huang, H., Gartner, G., Krisp, J. M., Raubal, M. & de Weghe, N. V. (2018), 'Location based services: ongoing evolution and research agenda', *Journal of Location Based Services* **12**(2), 63–93.

Hyafil, L. & Rivest, R. L. (1976), 'Constructing optimal binary decision trees is np-complete', *Information processing letters* **5**(1), 15–17.

Ishikawa, T. & Montello, D. R. (2006), 'Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places', *Cognitive Psychology* **52**(2), 93–129.

ISO (1985), 'Information processing - Documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts', *International Organization for Standardization, Geneva, CH* **ISO 5807:1985**.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013), *An introduction to statistical learning*, Springer.

Jansen-Osmann, P., Schmid, J. & Heil, M. (2007), 'Spatial knowledge of adults and children in a virtual environment: The role of environmental structure', *European Journal of Developmental Psychology* **4**(3), 251–272.

Jansen-Osmann, P. & Wiedenbauer, G. (2004), 'The representation of landmarks and routes in children and adults: A study in a virtual environment', *Journal of Environmental Psychology* **24**(3), 347 – 357.

Jonietz, D. (2016), From Space to Place: A Computational Model of Functional Place, PhD thesis, University of Augsburg, Faculty of Applied Computer Science.

Kamiński, B., Jakubczyk, M. & Szufel, P. (2018), 'A framework for sensitivity analysis of decision trees', *Central European Journal of Operations Research* **26**(1), 135–159.

Kass, G. V. (1980), 'An exploratory technique for investigating large quantities of categorical data', *Applied statistics* **29**(2), 119–127.

Kattenbeck, M. (2015), Empirically measuring salience of objects for use in pedestrian navigation, *in* 'Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems', ACM, pp. 3:1–3:10.

Kattenbeck, M. (2016), Empirically measuring salience of objects for use in pedestrian navigation, PhD thesis, University of Regensburg, Chair for information science.

Kattenbeck, M., Nuhn, E. & Timpf, S. (2018), Is salience robust? A heterogeneity analysis of survey ratings, *in* S. Winter, M. Sester & A. L. Griffin, eds, '10th International Conference on Geographic Information Science (GIScience 2018)', Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik.

Kim, S. & Lee, W. (2017), 'Does McNemar's test compare the sensitivities and specificities of two diagnostic tests?', *Statistical Methods in Medical Research* **26**(1), 142–154.

Klippel, A., Richter, K.-F. & Hansen, S. (2009), Cognitively ergonomic route directions, *in* H. A. Karimi, ed., 'Handbook of Research on Geoinformatics', IGI Global, pp. 230–238.

Klippel, A. & Winter, S. (2005), Structural salience of landmarks for route directions, *in* A. G. Cohn & D. M. Mark, eds, 'Spatial Information Theory. COSIT 2005', Vol. 3693 of *Lecture Notes in Computer Science*, Springer, pp. 347–362.

Kobsa, A., Koenemann, J. & Pohl, W. (2001), 'Personalised hypermedia presentation techniques for improving online customer relationships', *The Knowledge Engineering Review* **16**(2), 111–155.

Kohavi, R. (1995), A study of cross-validation and bootstrap for accuracy estimation and model selection, *in* 'International Joint Conference on Artificial Intelligence', pp. 1137–1145.

Kolios, A., Mytilinou, V., Lozano-Minguez, E. & Salonitis, K. (2016), 'A comparative study of multiple-criteria decision-making methods under stochastic inputs', *Energies* **9**(7), 566.

Krapp, A., Hidi, S. & Renninger, S. A. (2017), Interest, learning and developement, *in* R. Ann, H. Suzanne & K. Andreas, eds, 'The role of interest in learning and development', Psychology Press, pp. 3–26.

Krisp, J. M. (2016), 'Landmarks for Location-Based Services (LBS) in particular navigation and wayfinding', *KI-Künstliche Intelligenz* **31**(2), 199–201.

Krosnick, J. A. (1991), 'Response strategies for coping with the cognitive demands of attitude measures in surveys', *Applied Cognitive Psychology* **5**(3), 213–236.

Krosnick, J. A. & Fabrigar, L. R. (1997), Designing rating scales for effective measurement in surveys, *in* L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz & D. Trewin, eds, 'Survey Measurement and Process Quality', John Wiley & Sons, Inc., pp. 141–164.

Krumm, J., Davies, N. & Narayanaswami, C. (2008), 'User-generated content', *IEEE Pervasive Computing* **7**(4), 10–11.

Kuhn, M. & Johnson, K. (2013), *Applied predictive modeling*, Springer.

Kuipers, B. (1978), 'Modeling spatial knowledge', *Cognitive Science* **2**(2), 129–153.

Lin, N., Noe, D. & He, X. (2006), Tree-based methods and their applications, *in* H. Pham, ed., 'Springer Handbook of Engineering Statistics', Springer, pp. 551–570.

Loh, W.-Y. & Shih, Y.-S. (1997), 'Split selection methods for classification trees', *Statistica Sinica* **7**(4), 815–840.

Loomis, J. M., Golledge, R. G. & Klatzky, R. L. (2001), GPS-Based Navigation Systems for the Visually Impaired, *in* W. Barfield & T. Caudell, eds, 'Fundamentals of Wearable Computers and Augmented Reality', Lawrences Earlbaum Associates, Inc., pp. 444–461.

Lovelace, K. L., Hegarty, M. & Montello, D. R. (1999), Elements of good route directions in familiar and unfamiliar environments, *in* C. Freksa & D. M. Mark, eds, 'Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. COSIT 1999', Vol. 1661 of *Lecture Notes in Computer Science*, Springer, pp. 65–82.

LucidChart (2018), 'What is a flowchart'. Accessed September 2018.
URL: *https://www.lucidchart.com/pages/what-is-a-flowchart-tutorial*

Lynch, K. (1960), *The image of the city*, MIT press.

Maass, W. (1996), *Von visuellen Daten zu inkrementellen Wegbeschreibungen in dreidimensionalen Umgebungen: Das Modell eines kognitiven Agenten*, Infix.

Mack, A., Rock, I. et al. (1998), *Inattentional blindness*, MIT press Cambridge.

Mardani, A., Jusoh, A. & Zavadskas, E. K. (2015), 'Fuzzy multiple criteria decision-making techniques and applications–two decades review from 1994 to 2014', *Expert Systems with Applications* **42**(8), 4126–4148.

McGillivray, S., Murayama, K. & Castel, A. D. (2015), 'Thirst for knowledge: The effects of curiosity and interest on memory in younger and older adults.', *Psychology and Aging* **30**(4), 835–841.

Mendonça, L. F., Vieira, S. M. & Sousa, J. (2007), 'Decision tree search methods in fuzzy modeling and classification', *International Journal of Approximate Reasoning* **44**(2), 106–123.

Meng, L. (2005), 'Egocentric design of map-based mobile services', *The Cartographic Journal* **42**(1), 5–13.

Michon, P.-E. & Denis, M. (2001), When and why are visual landmarks used in giving directions?, *in* D. R. Montello, ed., 'Spatial Information Theory. COSIT 2001', Vol. 2205 of *Lecture Notes in Computer Science*, Springer, pp. 292–305.

Mingers, J. (1989), 'An empirical comparison of pruning methods for decision tree induction', *Machine Learning* **4**(2), 227–243.

Montello, D. R. (1998), A new framework for understanding the acquisition of spatial knowledge in large-scale environments, *in* E. M. J. & R. G. Golledge, eds, 'Spatial and temporal reasoning in geographic information systems', New York: Oxford University Press, pp. 143 – 154.

Montello, D. R. (2005), Navigation, *in* P. Shah & A. Miyake, eds, 'The Cambridge handbook of visuospatial thinking', Cambridge University Press, pp. 257–294.

Montello, D. R. & Sas, C. (2006), Human factors of wayfinding in navigation, *in* W. Karwowski, ed., 'International encyclopedia of ergonomics and human factors', Taylor & Francis, pp. 2003–2008.

Morio, J. (2011), 'Global and local sensitivity analysis methods for a physical system', *European Journal of Physics* **32**(6), 1577.

Mubayi, A. (2017), Computational modeling approaches linking health and social sciences: Sensitivity of social determinants on the patterns of health risk behaviors and diseases, *in* A. S. Rao Scinivasa, S. Pyne & C. Rao, eds, 'Handbook of statistics', Elsevier, pp. 249–304.

Murthy, S. K. (1998), 'Automatic construction of decision trees from data: A multi-disciplinary survey.', *Data Mining and Knowledge Discovery* **2**, 345–389.

Myler, H. R. (1998), *Fundamentals of engineering programming with C and Fortran*, Cambridge University Press.

Nisbet, R., Miner, G. & Elder, J. (2009), *Handbook of statistical analysis and data mining applications*, Elsevier.

Nothegger, C., Winter, S. & Raubal, M. (2004), 'Selection of salient features for route directions', *Spatial Cognition & Computation* **4**(2), 113–136.

Nuhn, E., Reinhardt, W. & Haske, B. (2012), Generation of landmarks from 3d city models and osm data, *in* J. Gensel, D. Josselin & D. Vandenbroucke, eds, 'Proceedings of the AGILE 2012 international conference on geographic information science', pp. 365–369.

Nuhn, E. & Timpf, S. (2016), A multidimensional model for personalized landmarks, *in* 'International Conference on Location Based Services, Vienna, Austria', Research Group Cartography, Vienna University of Technology, pp. 4–6.

Nuhn, E. & Timpf, S. (2017*a*), 'A multidimensional model for selecting personalised landmarks', *Journal of Location Based Services* **11**(3-4), 153–180.

Nuhn, E. & Timpf, S. (2017*b*), Personal dimensions of landmarks, *in* A. Bregt, T. Sarjakoski, R. van Lammeren & F. Rip, eds, 'Societal Geo-innovation: Selected papers of the 20th AGILE conference on Geographic Information Science', Springer, pp. 129–143.

Nuhn, E. & Timpf, S. (2017*c*), Towards personalized landmarks, *in* F. Paolo, B. Andrea & C. Eliseo, eds, 'Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory', Springer, pp. 101–103.

Nuhn, E. & Timpf, S. (2018), An overall framework for personalised landmark selection, *in* P. Kiefer, H. Huang, N. Van de Weghe & M. Raubal, eds, 'LBS 2018: 14th International Conference on Location Based Services', Springer, pp. 231–253.

O'keefe, J. & Nadel, L. (1978), *The hippocampus as a cognitive map*, Clarendon Press.

Ostasiewicz, W. (1982), 'A new approach to fuzzy programming', *Fuzzy Sets and Systems* **7**(2), 139 – 152.

Palmiero, M. & Piccardi, L. (2017), 'The role of emotional landmarks on topographical memory', *Frontiers in Psychology* **8**, 763.

Pannell, D. J. (1997), 'Sensitivity analysis: strategies, methods, concepts, examples', *Journal of Agricultural Economics* **16**, 139–152.

Passini, R. (1984), 'Spatial representations, a wayfinding perspective', *Journal of Environmental Psychology* **4**(2), 153 – 164.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A. & Cournapeau, D. (2011), 'Scikit-learn: Machine learning in python', *Journal of Machine Learning Research* **12**, 2825–2830.

Perrig, W., Wippich, W. & Perrig-Chiello, P. (1993), *Unbewusste Informationsverarbeitung*, Hans Huber.

Pertzov, Y., Dong, M. Y., Peich, M.-C. & Husain, M. (2012), 'Forgetting what was where: The fragility of object-location binding', *PLOS ONE* **7**(4), 1–12.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. & Wagener, T. (2016), 'Sensitivity analysis of environmental models: A systematic review with practical workflow', *Environmental Modelling & Software* **79**, 214 – 232.

Presson, C. C. & Montello, D. R. (1988), 'Points of reference in spatial cognition: Stalking the elusive landmark', *British Journal of Developmental Psychology* **6**(4), 378–381.

Quesnot, T. & Roche, S. (2014), 'Measure of landmark semantic salience through geosocial data streams', *ISPRS International Journal of Geo-Information* **4**(1), 1–31.

Quesnot, T. & Roche, S. (2015), Quantifying the significance of semantic landmarks in familiar and unfamiliar environments, *in* S. I. Fabrikant, M. Raubal, B. Michela, C. Davies, S. Freundschuh & S. Bell, eds, 'Spatial Information Theory. COSIT 2015', Vol. 9368 of *Lecture Notes in Computer Science*, Springer, pp. 468–489.

Quinlan, J. R. (1986), 'Induction of decision trees', *Machine Learning* **1**(1), 81–106.

Quinlan, J. R. (1990), 'Learning logical definitions from relations', *Machine Learning* **5**(3), 239–266.

Quinlan, J. R. (2014), *C4.5: programs for machine learning*, Morgan Kaufman Publisher.

Rao, H., Shi, X., Rodrigue, A. K., Feng, J., Xia, Y., Elhoseny, M., Yuan, X. & Gu, L. (2019), 'Feature selection based on artificial bee colony and gradient boosting decision tree', *Applied Soft Computing* **74**, 634 – 642.

Raubal, M. & Winter, S. (2002), Enriching wayfinding instructions with local landmarks, *in* M. J. Egenhofer & D. M. Mark, eds, 'Geographic Information Science. GIScience 2002', Vol. 2478 of *Lecture Notes in Computer Science*, Springer, pp. 243–259.

Reichenbacher, T. (2007), The concept of relevance in mobile maps, *in* G. Gartner, W. Cartwright & M. P. Peterson, eds, 'Location Based Services and TeleCartography', Springer, Berlin, Heidelberg, pp. 231–246.

Renniger, A. K. & Su, S. (2012), Interest and its development, *in* R. M. Ryan, ed., 'The Oxford handbook of human motivation', Oxford university press, pp. 167–190.

Rensink, R. A., O'Regan, J. K. & Clark, J. J. (1997), 'To see or not to see: The need for attention to perceive changes in scenes', *Psychological Science* **8**(5), 368–373.

Reyhana, Z., Fithriasari, K., Atok, M. & Iriawan, N. (2018), 'Linking twitter sentiment knowledge with infrastructure development', *Matematika* **34**(3), 91–102.

Richter, K.-F. (2007), A uniform handling of different landmark types in route directions, *in* S. Winter, M. Duckham, L. Kulik & B. Kuipers, eds, 'Spatial Information Theory. COSIT 2007', Vol. 4736 of *Lecture Notes in Computer Science*, Springer, pp. 373–389.

Richter, K.-F. (2008), *Context-Specific Route Directions: Generation of Cognitively Motivated Wayfinding Instructions*, DisKi 314 / SFB/TR 8 Monographs Volume 3. IOS Press.

Richter, K.-F. (2017), 'Identifying landmark candidates beyond toy examples - a critical discussion and some way forward', *KI-Künstliche Intelligenz* **31**(2), 135–139.

Richter, K.-F. & Duckham, M. (2008), Simplest instructions: Finding easy-to-describe routes for navigation, *in* T. J. Cova, H. J. Miller, K. Beard, A. U. Frank & M. F. Goodchild, eds, 'Geographic Information Science. GIScience 2008', Vol. 5266 of *Lecture Notes in Computer Science*, Springer, pp. 274–289.

Richter, K.-F. & Klippel, A. (2007), Before or after: Prepositions in spatially constrained systems, *in* T. Barkowsky, M. Knauff, G. Ligozat & D. R. Montello, eds, 'Spatial Cognition V Reasoning, Action, Interaction. Spatial Cognition 2006', Vol. 4387 of *Lecture Notes in Computer Science*, Springer, pp. 453–469.

Richter, K.-F. & Winter, S. (2011), Harvesting user-generated content for semantic spatial information: The case of landmarks in OpenStreetMap, *in* B. Hock, ed., 'Proceedings of the Surveying and Spatial Sciences Biennial Conference', Scion, pp. 75–86.

Richter, K.-F. & Winter, S. (2014), *Landmarks - GIScience for Intelligent Services*, Springer.

Riding, R. & Rayner, S. (1998), *Cognitive styles and learning strategies: Understanding style differences in learning and behavior*, David Fulton Publishers.

Rock, I., Linnett, C. M., Grant, P. & Mack, A. (1992), 'Perception without attention: Results of a new method', *Cognitive Psychology* **24**(4), 502–534.

Rokach, L. & Maimon, O. (2005), Decision trees, *in* 'Data mining and knowledge discovery handbook', Springer Science & Business Media, pp. 165–192.

Rokach, L. & Maimon, O. (2015), *Data mining with decision trees: theory and applications*, World scientific Publishing, Co.

Rosch, E. (1973), 'Natural categories', *Cognitive Psychology* **4**(3), 328 – 350.

Rosch, E. (1978), Principles of categorization, *in* M. Eric & L. Stephen, eds, 'Concepts: Core Readings', Cambridge, MA: MIT Press, pp. 189 – 206.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. (1976), 'Basic objects in natural categories', *Cognitive Psychology* **8**(3), 382–439.

Röser, F. (2015), The cognitive observer-based landmark-preference model - What is the ideal landmark position at an intersection?, PhD thesis, Justus-Liebig-University Giessen, Faculty 06 - Psychology and Sports Science.

Röser, F., Hamburger, K., Krumnack, A. & Knauff, M. (2012), 'The structural salience of landmarks: results from an on-line study and a virtual environment experiment', *Journal of Spatial Science* **57**(1), 37–50.

Rousell, A., Hahmann, S., Bakillah, M. & Mobasheri, A. (2015), Extraction of landmarks from OpenStreetMap for use in navigational instructions, *in* F. Bacao, M. Y. Santos & M. Painho, eds, 'Proceedings of the AGILE 2015 International Conference on Geographic Information Science', pp. 9–12.

Rousell, A. & Zipf, A. (2017), 'Towards a landmark-based pedestrian navigation service using osm data', *ISPRS International Journal of Geo-Information* **6**(3), 22.

Rüetschi, U.-J., Caduff, D., Schulz, F., Wolff, A. & Timpf, S. (2006), Routing by landmarks, *in* '6th Swiss Transport Research Conference', Institute for Economic Research, Università della Svizzera Italiana (USI), p. 15.

Ruotolo, F., Claessen, M. H. G. & van der Ham, I. J. M. (2018), 'Putting emotions in routes: the influence of emotionally laden landmarks on spatial memory', *Psychological Research* pp. 1–13.

Russell, S. J. & Norvig, P. (2016), *Artificial intelligence: a modern approach*, Pearson Education Limited.

Sa, I., Lehnert, C., English, A., McCool, C., Dayoub, F., Upcroft, B. & Perez, T. (2017), 'Peduncle detection of sweet pepper for autonomous crop harvesting - combined color and 3-d information', *IEEE Robotics and Automation Letters* **2**(2), 765–772.

Sadalla, E. K., Burroughs, W. J. & Staplin, L. J. (1980), 'Reference points in spatial cognition.', *Journal of Experimental Psychology: Human Learning and Memory* **6**(5), 516 – 528.

Sadeghian, P. & Kantardzic, M. (2008), 'The new generation of automatic landmark detection systems: Challenges and guidelines', *Spatial Cognition & Computation* **8**(3), 252–287.

Safavian, S. R. & Landgrebe, D. (1991), 'A survey of decision tree classifier methodology', *IEEE Transactions on Systems, Man, and Cybernetics* **21**(3), 660–674.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. & Tarantola, S. (2008), *Global sensitivity analysis. The primer*, John Wiley & Sons, Inc.

Sameer, A. & Bhushan, B. (2017), 'Effect of landmark type on route memory in unfamiliar homogenous environment', *Psychological Studies* **62**(2), 152–159.

Samuel, A. L. (1959), 'Some studies in machine learning using the game of checkers', *IBM Journal of Research and Development* **3**(3), 71–105.

Sandstrom, N. J., Kaufman, J. & Huettel, S. A. (1998), 'Males and females use different distal cues in a virtual environment navigation task', *Cognitive Brain Research* **6**(4), 351–360.

Sarjakoski, L. T., Koivula, T. & Sarjakoski, T. (2007), A knowledge-based map adaptation approach for mobile map services, *in* G. Gartner, W. Cartwright & M. P. Peterson, eds, 'Location Based Services and TeleCartography', Springer, pp. 247–264.

Sarjakoski, L. T. & Sarjakoski, T. (2008), User interfaces and adaptive maps, *in* S. Shekhar & H. Xiong, eds, 'Encyclopedia of GIS', Springer, pp. 1205–1212.

Schaffer, C. (1993), 'Selecting a classification method by cross-validation', *Machine Learning* **13**(1), 135–143.

Schmauks, D. (1998), 'Kognitive und semiotische Ressourcen für die Wegfindung', *Kognitionswissenschaft* **7**(3), 124–128.

Schober, M. F. (1998), Different kinds of conversational perspective-taking, *in* S. R. Fussel & R. J. Kreuz, eds, 'Social and cognitive psychological approaches to interpersonal communication', Lawrence Erlbaum Associates Publishers, pp. 145 – 174.

Schraw, G. & Lehman, S. (2001), 'Situational interest: A review of the literature and directions for future research', *Educational psychology review* **13**(1), 23–52.

Schroder, C. J., Mackaness, W. A. & Gittings, B. M. (2011), 'Giving the 'right' route directions: The requirements for pedestrian navigation systems', *Transactions in GIS* **15**(3), 419–438.

Scikit (2018), 'Documentation of skckit-learn 0.19.2'. Accessed August 2018.
URL: *http://scikit-learn.org/stable/documentation.html*

Scrivner-Limbaugh, A. L. T. (2015), Real-scene contextual cueing: why trees are better than cats, PhD thesis, University of Alabama, Department of Psychology.

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M. et al. (2016), 'Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information', *ISPRS international Journal of Geo-Information* **5**(5), 55.

Sensis (2017), 'Whereis routing service'. Accessed July 2017.
URL: *http://www.whereis.com/*

Shi, Y., Serdyukov, P., Hanjalic, A. & Larson, M. (2011), Personalized landmark recommendation based on geotags from photo sharing sites, *in* L. A. Adamic, R. A. Baeza-Yates & S. Counts, eds, 'Fifth International AAAI Conference on Weblogs and Social Media', AAAI Press, pp. 622–625.

Siegel, A. W. & White, S. H. (1975), 'The development of spatial representations of large-scale environments', *Advances in Child Development and Behavior* **10**, 9–55.

Silva, M. M., Groeger, J. A. & Bradshaw, M. F. (2006), 'Attention–memory interactions in scene perception', *Spatial Vision* **19**(1), 9–19.

Simons, D. J. (2000), 'Attentional capture and inattentional blindness', *Trends in Cognitive Sciences* **4**(4), 147–155.

Song, Y.-Y. & Lu, Y. (2015), 'Decision tree methods: applications for classification and prediction', *Shanghai Archives of Psychiatry* **27**(2), 130–136.

Sorrows, M. E. & Hirtle, S. C. (1999), The nature of landmarks for real and electronic spaces, *in* C. Freksa & D. M. Mark, eds, 'Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. COSIT 1999', Vol. 1661 of *Lecture Notes in Computer Science*, Springer, pp. 37–50.

Srinivasan, V. (2016), *The intelligent enterprise in the era of big data*, Wiley Online Library.

Steck, S. D. & Mallot, H. A. (2000), 'The role of global and local landmarks in virtual environment navigation', *Presence: Teleoperators and Virtual Environments* **9**(1), 69–83.

Stein, G., Chen, B., Wu, A. S. & Hua, K. A. (2005), Decision tree classifier for network intrusion detection with GA-based feature selection, *in* 'Proceedings of the 43rd annual Southeast regional conference-Volume 2', ACM, pp. 136–141.

Stone, M. (1974), 'Cross-validatory choice and assessment of statistical predictions', *Journal of the Royal Statistical Society. Series B (Methodological)* **36**(2), 111–147.

Sug, H. (2009), 'An effective sampling method for decision trees considering comprehensibility and accuracy', *W. Trans. on Comp* **8**(4), 631–640.

Sugumaran, V., Muralidharan, V. & Ramachandran, K. (2007), 'Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing', *Mechanical Systems and Signal Processing* **21**(2), 930 – 942.

Survey123 (2018), 'Survey123 for ArcGIS'. Accessed November 2018.
URL: *https://survey123.arcgis.com/*

Tan, P.-N., Steinbach, M. & Kumar, V. (2006), *Introduction to data mining*, Addison-Wesley Longman Publishing Co., Inc.

Tanaka, K. & Mizumoto, M. (1975), Fuzzy programs and their execution, *in* L. A. Zadeh, K.-S. Fu, K. Tanak & M. Shimura, eds, 'Fuzzy Sets and Their Applications to Cognitive and Decision Processes', Academic Press, Inc., pp. 41–76.

Tenbrink, T. & Winter, S. (2009), 'Variable granularity in route directions', *Spatial Cognition & Computation* **9**(1), 64–93.

Tezuka, T. & Tanaka, K. (2005), Landmark extraction: A web mining approach, *in* A. G. Cohn & D. M. Mark, eds, 'Spatial Information Theory. COSIT 2005', Vol. 3693 of *Lecture Notes in Computer Science*, Springer, pp. 379–396.

Thorndyke, P. W. (1980), Spatial cognition and reasoning, *in* J. H. Harvey, ed., 'Cognition, social behavior, and the environment', Lawrence Erlbaum Associates Publishers.

Tolman, E. C. (1948), 'Cognitive maps in rats and men.', *Psychological Review* **55**(4), 189–208.

Tom, A. & Denis, M. (2003), Referring to landmark or street information in route directions: what difference does it make?, *in* W. Kuhn, M. F. Worboys & S. Timpf, eds, 'Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003', Vol. 2825 of *Lecture Notes in Computer Science*, Springer, pp. 362–374.

Tomko, M. (2004), Case study-assessing spatial distribution of web resources for navigation services, *in* Y.-J. Kwon, A. Bouju & C. Claramung, eds, 'Proceedings of the 4th international conference on Web and Wireless Geographical Information Systems', Springer, pp. 90–104.

Tomko, M. & Winter, S. (2009), 'Pragmatic construction of destination descriptions for urban environments', *Spatial Cognition & Computation* **9**(1), 1–29.

Triantaphyllou, E. (2000), *Multi-criteria Decision Making Methods: A Comparative Study*, Springer Science & Business Media.

Ture, M., Tokatli, F. & Kurt, I. (2009), 'Using Kaplan-Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4. 5 and ID3) in determining recurrence-free survival of breast cancer patients', *Expert Systems with Applications* **36**(2), 2017–2026.

Tversky, B. & Lee, P. (1999), Pictorial and verbal tools for conveying routes, *in* C. Freksa & D. M. Mark, eds, 'Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. COSIT 1999', Vol. 1661 of *Lecture Notes in Computer Science*, Springer, pp. 752–752.

Vahle, F. (2014), *Die fabelhafte Geschichte von Anne Kaffeekanne*, Sauerländer Fischer Verlag GmbH, Frankfurt am Main.

Viaene, P., Vansteenkiste, P., Lenoir, M., De Wulf, A. & De Maeyer, P. (2016), 'Examining the validity of the total dwell time of eye fixations to identify landmarks in a building', *Journal of Eye Movement Research* **9**(3), 1–11.

Von Stülpnagel, R. & Steffens, M. C. (2013), 'Active route learning in virtual environments: disentangling movement control from intention, instruction specificity, and navigation control', *Psychological Research* **77**(5), 555–574.

Waller, D., Loomis, J. M., Golledge, R. G. & Beall, A. C. (2000), 'Place learning in humans: The role of distance and direction information', *Spatial Cognition and Computation* **2**(4), 333–354.

Wang, C., Chen, Y., Zheng, S. & Liao, H. (2019), 'Gender and age differences in using indoor maps for wayfinding in real environments', *ISPRS International Journal of Geo-Information* **8**(11), 1–20.

Wang, R. F. & Spelke, E. S. (2000), 'Updating egocentric representations in human navigation', *Cognition* **77**(3), 215 – 250.

Wang, Z., Wang, Y. & Srinivasan, R. S. (2018), 'A novel ensemble learning approach to support building energy use prediction', *Energy and Buildings* **159**, 109–122.

Ward, S. L., Newcombe, N. & Overton, W. F. (1986), 'Turn left at the church, or three miles north: A study of direction giving and sex differences', *Environment and Behavior* **18**(2), 192–213.

Wei*β*enberg, N., Gartmann, R. & Voisard, A. (2006), 'An ontology-based approach to personalized situation-aware mobile service supply', *GeoInformatica* **10**(1), 55–90.

Wei*β*enberg, N., Voisard, A. & Gartmann, R. (2004), Using ontologies in personalized mobile applications, *in* I. F. Cruz & D. Pfoser, eds, 'Proceedings of the 12th annual ACM international workshop on Geographic information systems', ACM, pp. 2–11.

Wender, K. F. (1998), 'Kontexteffekte und Routenwissen', *Kognitionswissenschaft* **7**(2), 68–74.

Wenig, N., Wenig, D., Ernst, S., Malaka, R., Hecht, B. & Schöning, J. (2017), Pharos: improving navigation instructions on smartwatches by including global landmarks, *in* M. Jones & M. Tscheligi, eds, 'Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services', ACM, p. 7.

Werner, S., Krieg-Brückner, B., Mallot, H. A., Schweizer, K. & Freksa, C. (1997), Spatial cognition: The role of landmark, route, and survey knowledge in human and robot navigation, *in* M. Jarke, K. Pasedach & K. Pohl, eds, 'Informatik '97 Informatik als Innovationsmotor: 27. Jahrestagung der Gesellschaft für Informatik Aachen', Springer, pp. 41–50.

Wiebrock, I. (2011), Zur kontextbasierten Visualisierung von Geodaten auf Basis von standardisierten Webdiensten, PhD thesis, University of the Bundeswehr Munich, Faculty of civil engineering and surveying.

Wiener, J. M., Büchner, S. J. & Hölscher, C. (2009), 'Taxonomy of human wayfinding tasks: A knowledge-based approach', *Spatial Cognition & Computation* **9**(2), 152–165.

Winter, S. (2003), Route adaptive selection of salient features, *in* W. Kuhn, M. F. Worboys & S. Timpf, eds, 'Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003', Vol. 2825 of *Lecture Notes in Computer Science*, Springer, pp. 349–361.

Winter, S., Duckham, M. & Robinson, M. (2009), 'Routing by landmarks', *Geoinformatics Magazine* **12**(7), 58–59.

Winter, S., Janowicz, K., Richter, K.-F. & Vasardani, M. (2012), 'Knowledge acquisition about places', *SIGSPATIAL Special* **4**(3), 20–21.

Winter, S., Raubal, M. & Nothegger, C. (2005), Focalizing measures of salience for wayfinding, *in* Z. A. Meng Liqui & R. Tumasch, eds, 'Map-based Mobile Services - Theories, Methods and Implementations', Springer, pp. 125–139.

Winter, S., Tomko, M., Elias, B. & Sester, M. (2008), 'Landmark hierarchies in context', *Environment and Planning B: Planning and Design* **35**(3), 381–398.

Wolfensberger, M. & Richter, K.-F. (2015), A mobile application for a user-generated collection of landmarks, *in* J. Gensel & M. Tomko, eds, 'Web and Wireless Geographical Information Systems. W2GIS 2015. Lecture Notes in Computer Science, vol 9080', Springer, pp. 3–19.

Wunderlich, A. & Gramann, K. (2018), Electrocortical evidence for long-term incidental spatial learning through modified navigation instructions, *in* S. H. Creem-Regehr, J. Schöning & A. Klippel, eds, 'Spatial Cognition XI - 11th International Conference, Spatial Cognition 2018', Vol. 11034 of *Lecture Notes in Computer Science*, Springer, pp. 261–278.

Yoon, K. P. & Hwang, C.-L. (1995), *Multiple attribute decision making: an introduction*, Sage publications.

Zadeh, L. (1965), 'Fuzzy sets', *Information and Control* **8**(3), 338 – 353.

Appendix A

Appendix

A.1 Tables

Table A.1: Parameter values for initial coarse grid-search CdTm.

Parameter	Value	Best Value
Criterion	gini, entropy	gini
Splitter	best, random	best
min_samples_split	[5, 10, ..., 50]	5
min_samples_leaf	[5, 10, ..., 50]	5
max_depth	[5, 10, ..., 50]	5
Average Accuracy [%]		76.19

Table A.2: Parameter values for finer grid-search CdTm.

Parameter	Value	Best Value
Criterion	gini, entropy	gini
Splitter	Best, Random	Random
min_samples_split	[2, 3, ..., 10]	2
min_samples_leaf	[1, 2, ..., 10]	1
max_depth	[1, 2, ..., 10]	4
Average Accuracy [%]		76.19

Table A.3: $\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$ for PspK and pInt ratings.

s_{PspK}	$s_{pInt(Shop)}$	$s_{pInt(Cult)}$	$s_{pInt(Hist)}$	$s_{pInt(Gast)}$	$\overline{s}_{vis}$	$\overline{s}_{sem}$	$\overline{s}_{str}$	
1	1	3	4	4	50.0	87.5	50.0	2
1	2	3	3	3	46.89	65.63	68.75	8
1	2	3	3	4	45.0	70.0	70.0	5
1	2	4	4	2	50.0	70.0	60.0	5
1	3	2	3	4	50.0	65.0	60.0	5
1	3	3	2	4	43.75	75.0	62.5	4
1	3	4	3	4	47.22	75.0	72.22	9
1	3	5	5	4	50.0	75.0	100.0	1
1	4	2	3	4	50.0	75.0	80.0	5
1	4	3	2	4	45.0	70.0	80.0	5
1	4	3	3	4	37.5	58.33	58.33	6
1	4	3	3	5	35.0	60.0	60.0	5
1	4	4	2	3	45.0	70.0	70.0	5
1	4	4	3	4	62.5	87.5	75.0	2
1	4	4	4	4	41.67	66.67	83.33	6
1	4	4	4	5	50.0	70.0	60.0	5
1	5	2	4	5	45.0	65.0	60.0	5
1	5	3	1	4	45.0	65.0	60.0	5
1	5	3	3	3	45.0	70.0	90.0	5
1	5	3	3	4	45.0	65.0	70.0	5
1	5	3	4	5	40.0	60.0	70.0	5
1	5	4	2	3	45.0	75.0	90.0	5
1	5	5	4	5	43.75	68.75	75.0	4
2	1	3	4	4	33.33	66.67	83.33	3
2	2	3	3	2	50.0	87.5	75.0	2
2	2	3	3	3	50.0	75.0	50.0	2
2	2	4	4	3	45.0	70.0	80.0	5
2	3	2	3	3	45.0	70.0	80.0	5
2	3	2	4	4	50.0	75.0	70.0	5
2	3	4	3	3	45.0	75.0	80.0	5
2	3	4	3	4	50.0	100.0	50.0	1
2	3	5	5	4	43.75	75.0	75.0	4
2	4	3	3	4	50.0	75.0	87.5	4
2	4	3	3	5	50.0	70.0	70.0	10

Table A.3: Continued $\overline{s}_{vis}$, $\overline{s}_{sem}$, and $\overline{s}_{str}$ for PspK and pInt ratings.

2	4	4	3	3	75.0	75.0	100.0	1
2	4	4	3	4	41.67	66.67	66.67	3
2	4	4	4	4	50.0	75.0	87.5	8
2	5	4	2	3	62.5	62.5	75.0	2
2	5	4	5	5	50.0	75.0	66.67	3
2	5	5	4	5	50.0	50.0	50.0	1
3	1	3	4	3	50.0	75.0	100.0	1
3	1	4	4	4	50.0	75.0	70.0	5
3	2	3	3	2	50.0	62.5	75.0	2
3	2	3	3	4	50.0	100.0	50.0	1
3	2	4	4	4	50.0	75.0	100.0	1
3	3	2	2	3	50.0	75.0	50.0	1
3	3	3	3	3	45.0	70.0	70.0	5
3	4	3	4	4	50.0	66.67	83.33	3
3	4	4	3	3	37.5	75.0	87.5	4
3	4	4	4	4	45.0	65.0	60.0	5
3	4	5	3	4	45.0	70.0	70.0	5
3	4	5	5	4	50.0	75.0	50.0	1
3	5	4	2	3	41.67	75.0	50.0	3
3	5	4	5	5	50.0	75.0	100.0	1
6	2	3	3	4	50.0	50.0	50.0	1
6	3	2	2	3	50.0	75.0	100.0	3
6	4	5	5	4	25.0	50.0	50.0	1
7	1	3	4	3	43.75	68.75	62.5	4
7	2	3	3	2	50.0	75.0	50.0	1
7	2	3	3	4	41.67	58.33	66.67	3
7	2	4	4	4	43.75	75.0	62.5	4
7	3	2	2	3	50.0	100.0	50.0	1
7	3	3	4	3	40.0	60.0	70.0	5
7	3	5	3	3	45.0	65.0	70.0	5
7	4	3	4	4	37.5	75.0	50.0	2
7	4	4	3	4	50.0	75.0	70.0	5
7	4	4	4	4	50.0	100.0	50.0	1
7	4	5	5	4	50.0	62.5	75.0	2

Table A.4: Average recalls of different flow charts with personal interests first.

$s_{PspK}(Intersection)$		Average Recall [%]
>3	≤ 3	
see Figure A.12	see Figure A.12	55.12
skip $max(s_{str})$	$s_{iLM} = s_{pInt} \rightarrow$ skip $max(s_{vis})$ and $max(s_{str})$	59.12

Table A.5: Results of sensitivity analysis of PwSm to s_{sem}.

ID	s_{sem}					SI
	0	25	50	75	100	
1	1.78	2.17	**2.56**	**2.94**	**3.33**	0.47
2	**2.56**	**2.56**	**2.56**	2.56	2.56	0
3	**2.56**	**2.56**	**2.56**	2.56	2.56	0

Table A.6: Results of sensitivity analysis of PwSm to s_{str}.

ID	s_{str}			SI
	0	50	100	
1	1.78	**2.56**	**3.33**	0.47
2	**2.56**	**2.56**	2.56	0
3	**2.56**	**2.56**	2.56	0

Table A.7: Example for sensitivity analysis of the PwSm to the landmark dimensions.

ID	s_{vis}	s_{sem}	s_{str}	$s_{iLM(shop)}$	$s_{iLM(cult)}$	$s_{iLM(hist)}$	$s_{iLM(gast)}$
1	0...100	100	100	1	1	1	1
2	100	100	100	1	1	1	1
3	100	100	100	1	1	1	1

Table A.8: Results of sensitivity analysis of PwSm to s_{vis} with $s_{vis} = s_{sem} = s_{str} = 100$.

ID	s_{vis}					SI
	0	25	50	75	100	
1	3.11	3.61	4.11	4.61	**5.11**	0.39
2	**5.11**	**5.11**	**5.11**	**5.11**	**5.11**	0
3	**5.11**	**5.11**	**5.11**	**5.11**	**5.11**	0
avgSI						0.13

Table A.9: Results of sensitivity analysis of PwPm to s_{sem}.

s_{vis}	ID		
	1	2	3
0	0	**42412775.26**	**42412775.26**
25	4908614.51	**42412775.26**	**42412775.26**
50	**42412775.26**	**42412775.26**	**42412775.26**
75	**149739425.95**	42412775.26	42412775.26
100	**366466647.78**	42412775.26	42412775.26
SI	1	0	0

Table A.10: Results of sensitivity analysis of PwPm to s_{str}.

s_{vis}	ID		
	1	2	3
0	0	**42412775.26**	**42412775.26**
50	**42412775.26**	**42412775.26**	**42412775.26**
100	**124671037.42**	42412775.26	42412775.26
SI	1	0	0

A.2 Figures

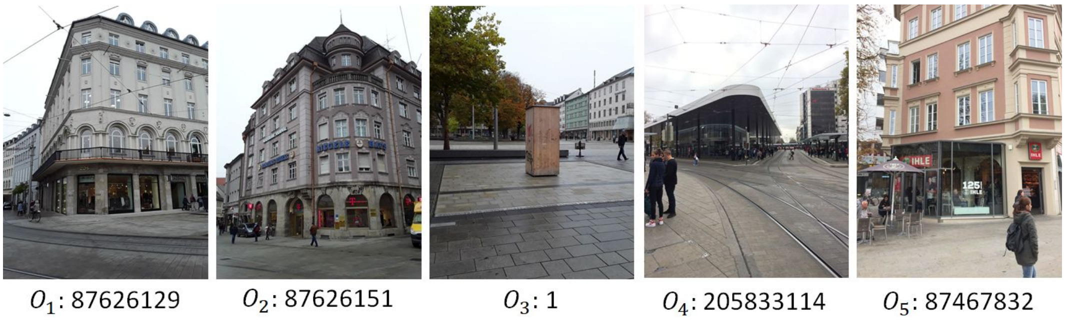


Figure A.1: Objects at decision point 0.

Figure A.2: Objects at decision point 1.

Figure A.3: Objects at decision point 2.

Figure A.4: Objects at decision point 3.

Figure A.5: Objects at decision point 4.

Figure A.6: Objects at decision point 5.

Figure A.7: Objects at decision point 6.

Figure A.8: Objects at decision point 7.

Figure A.9: Objects at decision point 8.

Figure A.10: Objects at decision point 9.

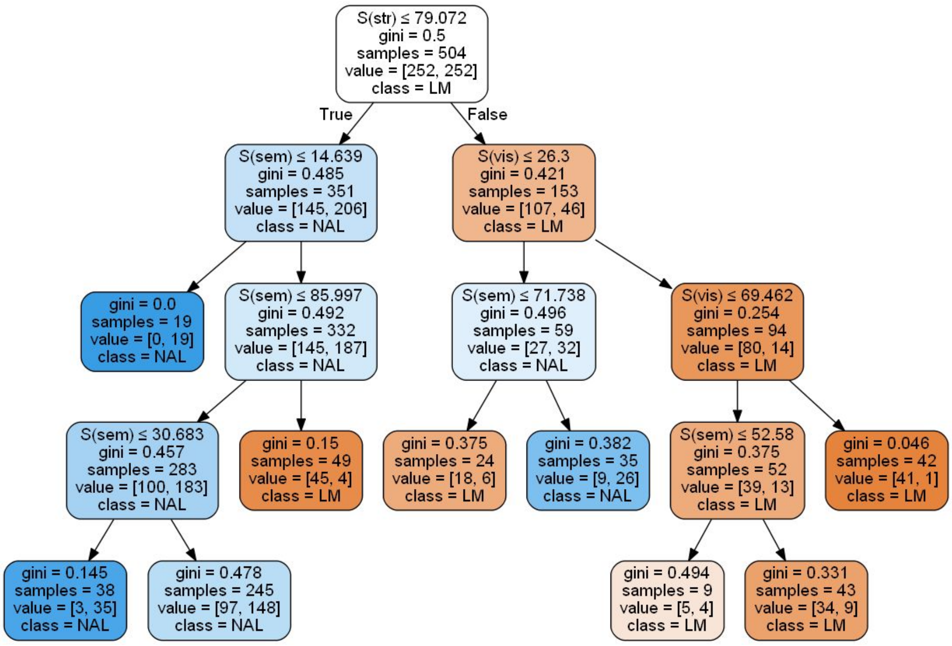


Figure A.11: Conventional Decision Tree Model.

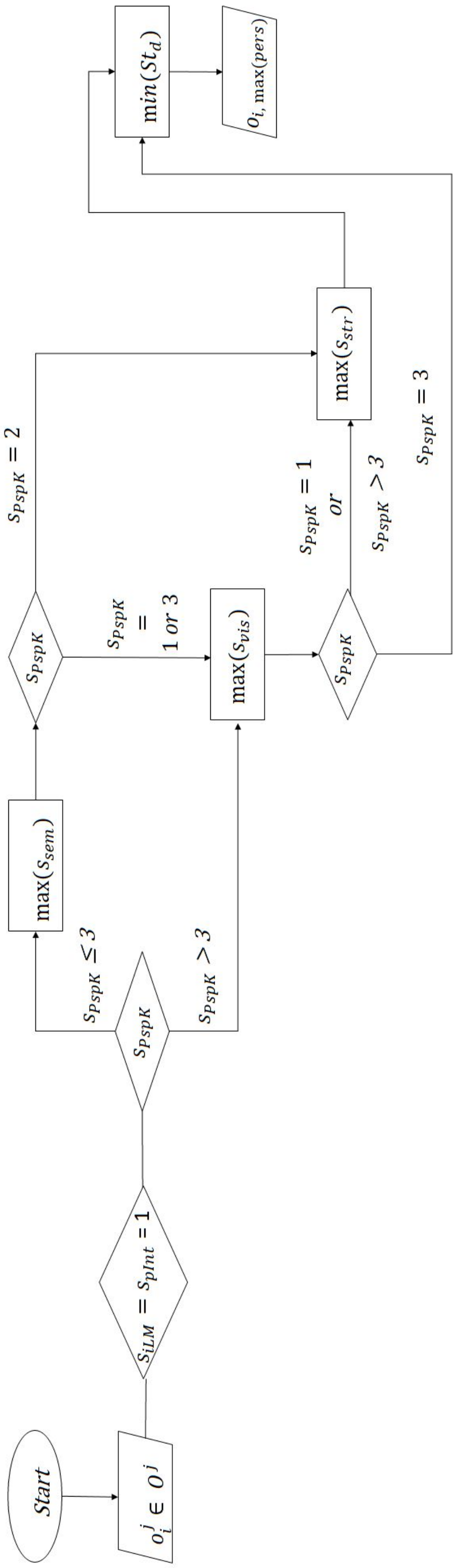


Figure A.12: Personalised Decision Flow Chart considering interest first.