Classification of Experience for Proactive In-Car Function Recommendations Based on Customer Usage Data

Christian Micus*
Technical University of Munich
Christian.Micus@tum.de

Daniel Homola*
BMW Group
Daniel.Homola@bmw.de

Markus Böhm
University of Applied Sciences Landshut
Markus.Boehm@haw-landshut.de

Helmut Krcmar
Technical University of Munich
Helmut.Krcmar@tum.de

Abstract

Automotive companies can use data from connected vehicles to enhance customer experience. Driver assistance functions have a low usage rate, and appropriate proactive function recommendations can improve both usage rate and customer experience. Qualitative studies often drive the development, and functions are recommended using a rule-based system. We provide a patented machine learning-based classification concept to make intelligent function recommendations based on customer usage. Therefore, we classify customer experience based on the driving context. We defined how to create an experience label for a function activation context and evaluated the approach using 716,000 function activations collected from the customer fleet data by an automotive manufacturer. To improve the quality of the binary classification model, we defined geospatial key performance indicators that provide quantifiable measures for the performance of a function on a road section. Our results reveal that the novel classification concept is a viable solution for car function recommendations.

Keywords: Connected Car, Proactive Services, Binary Classification, Ensemble Models, Geospatial Data

1. Introduction

Customers are finding it increasingly difficult to operate functions in complex products. At present, products are being developed with increasing functions intended to capture customers' interest and demonstrate the products' excellence. The development of new functions is associated with high costs for the manufacturer. Hence, manufacturers must judge which products and functions they want to develop are within their budget. A crucial point in developing new functions is that the functions must have customer value. With rising competition, this factor can lead to recurring product purchases. However, the challenge is that many customers often do not engage with the new functions and attempt to use them. Similarly, trying out these functions may not be effective because they cannot be used or do not provide a satisfactory experience in a particular situation. Therefore, the customer is annoyed by the new function and dislikes it.

Hence, proactive function recommendations are becoming increasingly crucial for companies. This study provides a machine learning-based concept for recommending a function to the customer at the right time. This is the challenge with proactive function recommendations because they must be made at the appropriate time to result in a positive experience for the customer. We conducted a concrete case study in the automotive industry to present a context classification approach for delivering proactive function recommendations. For this case study, we selected a driver assistance function. It has a usage rate of 34% (i.e., it is activated at least once during a drive), and its activation by the customer depends on the driving context. In this study, a patented machine learning model was developed to predict proactive function recommendations, and the model results were compared with those of an existing system that recommends a function using predetermined conditions (Homola & Micus, 2023). Furthermore, this paper describes the advantages of a machine learning-based model and presents an approach toward developing more business models.

1.1 Proactive services

Proactive services and communication enable a driver to receive recommendations when a function is safely

* Christian Micus and Daniel Homola contributed approximately equally to this paper.

URI: https://hdl.handle.net/10125/106484
978-0-9981331-7-1

applicable and stable functionality is expected. Such services ensure that the function meets the user's expectations and that the experience increases user satisfaction with the vehicle. Consequently, increased customer satisfaction and demand for relevant functions in future vehicle purchases are expected. Original equipment manufacturers who focus on connected services may improve their market position. According to a 2020 study on connected car services and technologies (NTT-Data, 2020), 85% of survey participants see an added value in proactive services (as the most valuable connected car functions). Functions are better understood in real driving situations (as opposed to explaining the function during a vehicle handover). The driver can directly try out the appropriate assistance system in an actual driving situation. Proactive function recommendation and their simple confirmation improve the customer experience since the customer no longer has to think about whether and when he can use a particular function. Similarly, proactive function recommendations are highly relevant to deploying and using new business models, such as subscription models for using functions or pay-per-use functions. Proactive recommender systems must determine which function should be recommended and when a function recommendation should be made. Thus, the design of proactive recommender systems in an automotive environment should be investigated (Bader, 2013).

1.2 Importance of customer usage data

In an environment of increasing volatility of external influences, shorter product life cycles, increasing global competition, and exponential technological advancements, traditional manufacturers face disruptive business changes (Huang et al., 2022). Product development is tasked with responding to customer needs, adapting to environmental changes, and providing customer satisfaction (Einizadeh & Kasraei, 2021). One of the key approaches to ensuring customer satisfaction and improving the competitive advantage is to increase customer integration in the development process (Micus et al., 2022). This customer-centric development process reduces uncertainty in customer demands (Wang et al., 2021), allowing for the targeted allocation of resources and the development of concepts that accurately meet customer needs and have a high usage frequency (Koushik & Mehl, 2015). With increasingly connected vehicles, automotive manufacturers can access live vehicle data from the customer fleet (Abbasi et al., 2016). Data analytics methods can be used to analyze vehicle data from an entire customer vehicle fleet in real time and to make product improvement decision. Information technology and data processing are the key to capturing business events and reacting swiftly to changes, resulting in drastic changes in how businesses use data in decision-making (Abbasi et al., 2016; Jones, 2019). Companies strive to integrate, prepare, and analyze data in real-time (Micus et al., 2023).

1.3 Recommendations based on customer usage data

Product development of proactive services is often driven by customer surveys and lead-user workshops in the automotive industry. However, the results reflect the intention to use rather than the actual use of the customer (see literature on the technology acceptance model (TAM)) (Davis, 1985). Consequently, product development is driven primarily by the intention to use rather than the actual use of a function (Enkel et al., 2005). The same applies to proactive function recommendations. For example, the function recommendations can be driven by fixed conditions that experts predetermine via surveys or workshops. Nevertheless, such conditions are strictly chosen to avoid inadequate function recommendations and guarantee proper function operation. Rather than having fixed conditions, incorporating customer usage data to derive the appropriate function recommendation scenarios can provide a remarkable advantage. The approach of fixed conditions primarily uses the current context of the vehicle. However, more information is hidden in the customer usage data than in the current context when GPS-related features are considered. Suppose many drivers do not use a particular driver assistance function in the current road section. Even if the current context of the vehicle is suitable for a function recommendation, that function may not be recommended after considering the customer usage data and determining that the function is not suitable for use for the given road section. The road section provides an abstraction of the precise geographic location, and we can learn more from such features on a higher abstraction level. A concept of a proactive service based on experience classification was designed to leverage the usage data of customer fleets collected by automotive manufacturers. The data includes past activation contexts of a function and the experience labels corresponding to those activation contexts. A supervised learning model can be trained for experience classification using this concept and the data. This study answers the following questions: “How can customer usage data be used to recommend functions proactively?” “What are the beneficial geospatial performance measures that can be created based on customer usage data?" and "What are the benefits of using customer usage data for ML-based function recommendations?"

2. State of the art

Recently, vehicle proactive recommender systems have attracted research interest (Guinea et al., 2020). Studies have shown that proactive vehicle recommendation systems are perceived as helpful and supportive but not obstructive or distracting while driving (Bader et al., 2011). Proactivity in technical systems is defined as the autonomous, anticipatory behavior initiated by the system, and it acts in anticipation of a future situation rather than merely reacting to the present situation (Nothdurft et al., 2015). The system becomes active without a conscious human-driven impulse. When an onboard vehicle computer initiates a dialog to recommend a function to the driver, the system is called proactive. For example, the proactive recommendation can be a suggestion for activating an assistance function. Proactive recommendations are predictive, i.e., based on predictions and estimates of future events and user needs. Moreover, they are based on implicit rather than explicit information (Nothdurft et al., 2015). They identify needs, prepare users for future events, and initiate practical or necessary actions early (Nothdurft et al., 2015). The challenge for the recommender systems is mapping the complex human decision-making processes. In particular, incorrect activation requests for automated driving functions have serious consequences (Guinea et al., 2020). Proactive function recommendations are designed to show the currently suitable custom functions. Thus, the customer learns about a customer value function, which convinces the customer of his product and achieves a higher usage rate. Two types of in-vehicle proactive function recommendations exist to determine the right moment to make a recommendation:

Recommendations based on fixed conditions: Customer surveys and lead-user workshops often drive recommendations. The results determine fixed conditions in which a function recommendation should be triggered. Furthermore, the recommendation can only be served if all predetermined conditions are fulfilled. When all conditions for the function are fulfilled, a positive label is assigned to the context, and the function can be recommended. This approach of recommendations based on fixed conditions can be evaluated similarly to a classification model. It is suitable for function recommendations or other suggestions with low driving complexity, such as a suggestion to take a break after 3 hours of driving.

Recommendations based on an ML model: Recommender systems are ML models that make data-driven decisions. They can integrate various features into decision-making (Cami et al., 2019). The ML model itself decides which features are essential for making a decision. The recommendation systems use context information of the user (e.g., location, time, and companion). This research field has become increasingly important for many researchers and practitioners. The application domain is diversified—for example, film, music, and e-commerce industries (Panniello et al., 2014). Automobile industry applications include intelligent infotainment, chassis, and cruise control systems (Bruss & Pfalzgraf, 2016; Lefèvre et al., 2015). There are currently no concepts for proactive function recommendations of driver assistance systems.

3. Methodology

We designed a patented concept of a proactive recommendation based on a context classification to leverage the customer usage data collected from car fleets by automotive manufacturers (Homola & Micus, 2023). The cars in the customer fleets report events of activation and deactivation of functions via available sensors and communicate using networks. Then, the event-based connected car data are processed into a structure where each data row represents the activation context of a function (car context, environment context, and driver context), duration of usage, and label of the experience that followed the given activation context. Subsequently, a supervised learning model was trained for the context classification of a potential function experience. Figure 1 shows an example of activation and deactivation points for using a function. The experience duration was computed using these two points, and the label for the respective activation context based on the duration was decided. The label is motivated by the assumption that if a function stays activated for at least a minimum duration (positive label), most likely, the driver appreciates a supportive experience of the function. In contrast, a short activation duration (negative label) may indicate the driver is dissatisfied with the function in the respective context.

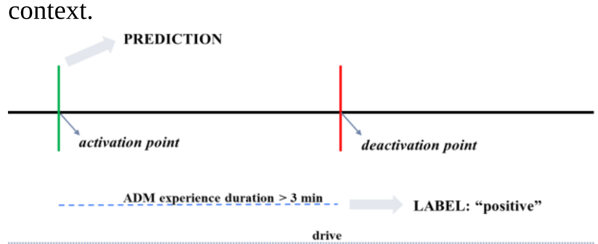

Figure 1: Classification concept and labeling.

Therefore, this study framed the problem of recommending the function at the right moment. The recommendation is based on context classification — the function is recommended when the context will lead to a positive function experience (i.e., the classification resulted in a positive experience label). This section further explains the classification solution, the classification model, and relevant metrics used to measure the performance of the classification model.

3.1 Concept of recommendation via binary classification

This proactive function recommendation problem is formulated as a binary classification: "Will the current context lead to a negative or positive experience for this function?". In particular, a positive experience can be associated with a supportive experience for the driver, and the function is qualified as a candidate for the recommendation. Thus, function recommendation can be based on the recommendation model's classification of the context information. The context information represents the current driving situation, and the recommendation model is a classification model. The dataset contains function activation contexts and assigns experience labels to the activation contexts. Each activation context can lead to a unique experience. Figure 2 illustrates the concept of classifying the customer experience. A car is on the road in a specific driving context, and this context is input to the machine learning model. The model then classifies if this driving context would lead to a positive or negative experience for the driver regarding the assistance function. If a positive experience is predicted, the function can be recommended to the driver because it is expected to provide a supportive experience. The machine learning model learns from past customer behavior: it learns from the past activations of a specific function.

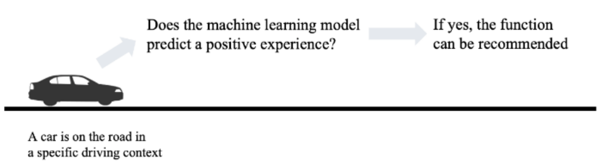

Figure 2: Illustration of driving context classification.

3.2 Ensemble classification models

A classification task can be solved using various classification models. An ensemble model was chosen because of the complexity of the task and the variety of features. Ensemble models use a set of estimators instead of a single base estimator to classify an outcome (Rokach, 2010). The base estimators of an ensemble are combined to make a prediction; thus, the ensemble usually performs better than a single estimator. Moreover, ensemble models tend to outperform single models in data science competitions (Bojer & Meldgaard, 2021). There are two families of ensemble methods: Averaging and boosting methods. In the averaging methods, base estimators are created independently, and their predictions are averaged. In the boosting methods, base estimators are built sequentially, and a strong ensemble is produced by combining several weak models. Among the averaging methods, we tested RandomForestClassifier and BaggingClassifier. Among the boosting methods, we tested AdaBoostClassifier, GradientBoosting Classifier, LGBMClassifier, XGBClassifier, and CatBoostClassifier. Their implementations in scikit-learn, lightgbm, xgboost, and catboost packages were used (Bojer & Meldgaard, 2021). The RandomForestClassifier provided the best results. Therefore, it was used in this study. The threshold for dichotomizing predictions of the random forest is 0.5 (an instance is classified as positive if predicted probability is greater than 0.5 and negative otherwise).

3.3. Data description

First, a dataset containing all activation and deactivation contexts of the function was obtained. Then, we assigned a label to each activation context based on the activation duration of the function in question (derived by subtracting the activation time from the deactivation time of the corresponding data points). This helped us immensely because we did not have to label the data manually. The final dataset for the machine learning problem contains the activation context feature and created label. In all, 21 features related to the function activation context are available and can be classified into three categories: the car context (e.g., driving speed), weather (e.g., outdoor temperature), or environment (e.g., number of objects around the car). Each feature has different importance for the models, and the classification accuracy, and some features must be further processed to train a model that uses them. In this study, we defined two labels (positive and negative) that were used for the supervised training of the models. The positive labels led to a supportive experience for the driver (i.e., activation duration exceeded a threshold duration), and the function was considered a candidate for the recommendation. The negative labels led to a somewhat discouraging experience for the driver (i.e., activation was below a threshold). This was motivated by 15 expert interviews, conducted specifically for this function, wherein we evaluated multiple intervals of activation durations. We found that intervals longer than 3 min lead to a positive, encouraging experience of our automotive function. A duration of 3 min was set as the business objective for this function. The threshold must be adapted to the objectives of each function and may vary among functions. Furthermore, data analysis of the experience duration was also a factor in deciding the threshold (see Figure 1). The analysis confirmed that the 3 min threshold of function usage is a reasonable split between positive and negative experiences. Our dataset contains 530351 negative experiences (74%) and 186608 positive experiences (26%) across Germany (see Figure 3). This indicates the presence of an imbalanced dataset, which, if addressed, may lead to potentially improved results for the proposed model.

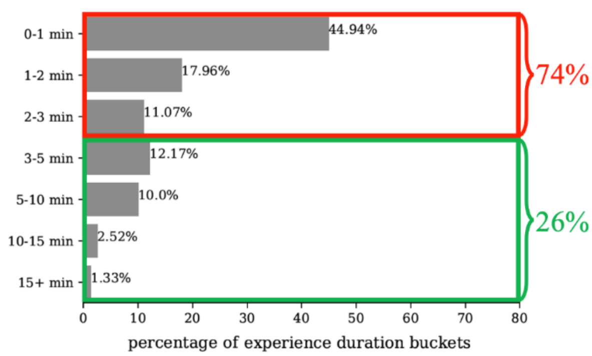

Figure 3: Buckets of function experience duration.

3.4. Data preparation

Many methods are usually used during feature engineering/data preparation (Micci-Barreca, 2001). We used multiple methods:

1. We solved the missing data problem by inputting an arbitrary number or string or removing the specific row if many values were missing.
2. We discretized some continuous variables to extract higher-level information.
3. We ordinally encoded categorical variables into integer arrays.
4. Some outliers that were distant from other values of the same feature were removed. To identify the outliers, we computed the interquartile range and filtered out the extreme outliers that lie more than thrice in the interquartile range below the first quartile or above the third quartile.

3.5. Geospatial key performance indicators

To create additional features, we aggregated the feature values. Such aggregations are also key performance indicators (KPIs; i.e., quantifiable performance measures). In this study, a geospatial key performance indicator is a numerical value associated with a driver assistance function and with the performance of this function on a specific road section (Parmenter, 2015). A road section is a defined part of a road to which we can assign certain properties e.g., how our customers drive the vehicle and which functions are activated on this section. Adding the KPIs to our model may produce training higher-quality datasets as road section identifiers provide a reasonable abstraction of the precise geographic location of the vehicle (many geographic coordinates belong to a single road section). Thus, the function recommendation can be further improved.
Geospatial KPIs are as follows (see Table 1):

- Activation Ratio: How often is the function activated on this road section?
- Deactivation Ratio: How often is the function deactivated on this road section?
- Function Usage Ratio: How often do customers traverse the road section with the activated function?
- Positive Ratio: How often do customers have a positive experience of the function if they activate the function on this road section? (target encoding)

We adopted the following definitions:

- activation_count: total number of function activations on a road section
- deactivation_count: total number of function deactivations on a road section
- function_traversal_count: total number of traversals of a road section with an activated function
- general_traversal_count: total number of traversals of a road section
- positive_experience_count: total number of function activations associated with a positive experience.

For the positive ratio, we encoded the target into a statistical indicator for each categorical feature (Micci-Barreca, 2001). In this target encoding concept, we replaced a categorical value with the mean of the target variable. Negative and positive experience labels were used to compute the mean target value per road section. Only the training set must be used to compute the mean target value; else, target information may leak from the test data evaluating the classifier. Extra regularization is necessary to compute target encodings with the training set (e.g., splitting the training set into multiple folds, computing the target encoding in one fold based on the other folds, and finally, applying the smoothing techniques).

Table 1: Formulas for the key performance indicators (KPIs).

KPI	Formula
Activation Ratio	$\frac{\text{activation\_count}}{\text{general\_traversal\_count}}$
Deactivation Ratio	$\frac{\text{deactivation\_count}}{\text{function\_traversal\_count}}$
Function Usage Ratio	$\frac{\text{function\_traversal\_count}}{\text{general\_traversal\_count}}$
Positive Ratio	$\frac{\text{positive\_experience\_count}}{\text{activation\_count}}$

3.6. Class imbalance

A class imbalance was observed in the dataset as 74% (530351 experiences) were labeled as negative experiences, and 26% (186608 experiences) were labeled as positive experiences. Therefore, methods were applied to resolve this issue. The relevant class can be undersampled or oversampled (Kotsiantis et al., 2006). The undersampling process decreases the size of the majority class. Undersampling can be a good choice when we have much data, but the disadvantage is that we might eliminate valuable information, leading to underfitting and poor generalization. The oversampling process involves increasing the size of the minority class. Oversampling can likely cause overfitting because the existing examples are copied (i.e., the minority class is replicated). We adopted a random oversampling method where the samples of the minority class are randomly duplicated. The amount of samples of negative class was matched by oversampling the samples of positive class. The classes are balanced only in the training set.

3.7. Metrics

Models for binary classification can be evaluated using various metrics (Canbek et al., 2017; Raschka, 2014). We evaluated the ensemble models primarily with three base metrics: accuracy, sensitivity, and specificity. Accuracy is not the best metric to evaluate imbalanced datasets as it can be misleading. Therefore, specificity and sensitivity were considered to cover the negative and positive experience classifications in two separate metrics. Additional metrics, such as precision (positive predictive value) and negative predictive value, were also computed.

Sensitivity (TPR = true positive rate, i.e., recall) measures the fraction of correctly identified positives, whereas specificity (TNR = true negative rate) measures the fraction of correctly identified negatives. Specificity covers true negatives, which are important because we do not want to recommend a function in a situation that would lead to a negative experience (i.e., a recommendation that can be associated with a discouraging experience). In making a trade-off between recommending a function, we would rather not recommend the function at a moment that would lead to a positive experience than recommend the function in a moment that would lead to a negative experience. This means we prefer to achieve as many true negatives as possible and as few false positives as possible. The metrics and their formulas are listed in

Table 2: Metrics.

These formulas are based on the following outcomes (classification results):

- True Positive (TP): The classifier correctly predicts the positive class.
- True Negative (TN): The classifier correctly predicts the negative class.
- False Positive (FP): The classifier incorrectly predicts the positive class.
- False Negative (FN): The classifier incorrectly predicts the negative class.

Table 2: Metrics.

Metric	Formula
Accuracy	$\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$
Sensitivity	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$
Specificity	$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$
Precision	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$
Negative predictive value	$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}$

4. Implementation of the data preparation steps

Before training the model with input features, we performed the following data preparation steps:

1. Feature engineering steps (imputation, discretization, ordinal encoding, and outlier handling)
2. Computation of simple geospatial KPIs (i.e., the activation ratio, deactivation ratio, and function usage ratio)
3. Computation of the target encoded geospatial KPI (i.e., the positive ratio)
4. Downsampling/oversampling the target classes

The combination of various steps to preprocess data and the model performance is presented in Table 3 (the explanations of why such steps were considered are provided in square brackets).

Table 3: Results after the data preparation steps.

RandomForestClassifier	**Ac[1]**	**Se[2]**	**Sp[3]**
Results after applying general feature engineering steps (1).	0.74	0.04	0,98
Results after applying general feature engineering steps (1) and computing simple geospatial KPIs (2). [Impact of simple KPIs].	0.74	0,06	0.98
Results after applying general feature engineering steps (1) and oversampling the minority class (4).	0.73	0.10	0.96
Results after applying general feature engineering steps (1), computing simple geospatial KPIs (2), and oversampling the minority class (4). [Combination of simple KPIs and oversampling].	0.74	0.12	0.95
Results after applying general feature engineering steps (1), computing simple geospatial KPIs (2), computing target encoded KPI (3), and oversampling the minority class (4). [Combination of simple KPIs, target encoded KPI and oversampling].	0.73	0.26	0.89

5. Comparison of fixed conditions and ML models

To validate the performance of the machine learning model, we compared it with the existing system's performance with fixed conditions. For validation, we analyzed the potential of the concept and compared the following three approaches for classifying experiences based on the context:

- Fixed conditions are determined for six features (the recommendation is made only if all the predetermined conditions proposed by experts are met).
- The basic ML model uses the same six features as the fixed conditions model.
- The complete ML model had 21 features, and it was obtained after the fourth data preparation step as detailed in the previous section.

The performances of the three approaches were compared in accuracy, sensitivity, and specificity. The most important metric in this study was specificity because we wished to avoid negative experiences as much as possible. The vehicle customer value is a recommendation of the function when it works well. Thus, it is more important for us to correctly identify the negative experiences (specificity) than correctly identify positive experiences (sensitivity). We would rather miss out on function recommendations that would lead to a positive experience than recommend in a moment that would lead to a negative experience. The data were split in training and test set, training set containing 80% of original samples and test set containing 20% of original samples. The reported classification results are on the test set.

Table 4 summarizes the performances of the different approaches.

Table 4: Classification results.

Approach	**Ac[1]**	**Se[2]**	**Sp[3]**
Fixed conditions (6 features)	0.56	0.35	0.64
Basic ML model (6 features)	0.69	0.15	0.89
Complete ML model (21 features)	0.73	0.26	0.89

6. Discussion

The following two sections summarize the findings related to the research questions. The results obtained with the ML model were compared with those obtained under fixed conditions, and the potential of the ML model was validated. We showed that creating a classification concept based on machine learning and a classification model for proactive function recommendations is feasible and valuable. We also showed that additional feature engineering and computing extra geospatial features can improve model quality. This section presents the interpretation and implications of the results and the limitations of this study.

[1] Accuracy

[2] Sensitivity

3 Specificity

6.1. Proactive function recommendations based on customer usage data

Comparing the classification approaches shows that building an ML-based model using customer usage data is possible to recommend functions successfully. A function experience can be modeled based on past activation contexts, and the context classification can be used as a basis for making a recommendation.

We found that processing the data with various additional data preparation steps can improve the model performance. The model accuracy improved after implementing different data preparation steps. We consider the models that are more balanced and that have higher sensitivity as better models than those that classify almost all activation contexts as negative experiences [very high specificity (greater than 0.9) and very low sensitivity (lower than 0.1)]. Thus, the data preparation steps are valuable, and it is essential to balance the target classes to achieve more balanced results.

We created geospatial performance measures by aggregating customer usage data on map data. These geospatial KPIs can also be used as input features to improve and balance the ML model. Moreover, we can use them to enhance the recommendation process, for example, by combining an ML model with a threshold for a positive ratio at which a given function recommendation should be triggered.

We also reported the classification results for recommendations based on fixed conditions and recommendations based on machine learning models (i.e., basic and complete ML models, respectively). The comparison shows that a classification model based on customer usage data can perform equally or even better than the model based on fixed conditions. Both basic and complete models perform 25% better than the model based on fixed conditions in correctly identifying the negatives (specificity) at the expense of worse performance in correctly identifying the positives (sensitivity). The complete model is much more precise than the basic model in classifying positives (precision of the complete model = 0.46; precision of the basic model = 0.26), despite correctly identifying 9% less positives (sensitivity of the complete model = 0.26; sensitivity of fixed conditions = 0.35). The complete model is better than the basic model in terms of specificity, negative predictive value, and accuracy; it is worse only in terms of sensitivity.

Overall, we achieved a better result with the random forest model (complete ML model with 21 features) than with the basic model.

6.2. Advantages of ML-based proactive function recommendations

Recommendations based on machine learning are better than those based on fixed conditions from the following viewpoints:

Performance: Learning from more features results in better performance. The complete model used 16 features more than that in the basic model, and hence, it performed 25% better than the basic model in terms of not recommending contexts that led to a negative experience. It was also more precise in classifying both positive and negative labels.

Flexibility: The ML model is more flexible. As components change over time, the system can change how the function is used. The ML model detects changes in usage and adapts to new circumstances through model updates. In doing so, the system determines the essential features for proactive function recommendations. Using the ML model, the system can continuously improve or avoid degradation. The system reacts to the changes in not only customer usage behavior but also the environment. The system notices the changes in road layout or road works immediately. The function usage in these road sections changes, and system failure hotspots where no proactive function recommendations should be offered are detected.

Robustness: The complete model is capable of more robust decision-making because its recommendations are based on more features in the machine learning model. Defining fixed conditions out of the features of the complete model would be difficult.

Customer orientation: Customer usage drives the recommendations, not just the qualitative intention to use (conditions). Thus, the model recommendations do not reflect the intention to use the function, as in the case of TAM research (see TAM literature), but the actual use of the system users [19].

Personalization potential: The importance of personalization is increasing immensely in all fields since customers expect a product that is optimally tailored to their needs. In-depth machine learning-based personalization and learning from user feedback per context will be possible for the machine learning model in the future. Analysis of customer usage data can help generate vehicle usage profiles of customers. Vehicle usage can be depicted by representing three dimensions: mobility behavior, usage of driver assistance systems, and driving dynamics. Using such features to represent the usage clusters of the vehicle profiles, customers' needs can be addressed accurately, and proactive function recommendations can be further personalized.

6.3. Limitations

The data mainly limits the analysis reported herein. This study focuses on a single driver assistance function, and only data from this function was processed. Nevertheless, the classification concept can be used for other functions as well. Furthermore, the data were obtained from one country (the whole of Germany) for one month (September 2020). This limits the generalization of the model owing to a country bias. Although it is desirable to have data from different countries over periods longer than one month, the data was sufficient to validate the concept and models. We used the data from Germany only as a validation use case for the designed approach. Based on the available data, the models can be trained for any country in a real production setup. Moreover, the models can be retrained over time based on new data.

7. Contribution

The results of this study contribute to theoretical and practical research. In theory, a novel patented proactive function recommendation concept was designed based on context classification and customer usage data (Homola & Micus, 2023). We processed customer usage data and developed an ML model for function recommendations. These customer usage data were enriched by linking them to map information, and additional geospatial road section KPIs were defined and provided to the ML model as input features. In the current state of the art, there is still no concept of using geospatial performance indicators from customer usage data for proactive function recommendations. Similarly, the literature has no comparable concept for customer usage-based proactive function recommendations in the vehicle.

The results of the ensemble model demonstrate that the experience classification of the activation context is a viable solution for driver assistance function recommendations. The recommendation approach achieved better results than the current approach based on fixed conditions (i.e., rule-based recommendations). The new concept can minimize the uncertainty of the right moment for proactive function recommendations. This proves that the classification approach can be used in the automotive industry.

We elucidated the benefits of the ML model and explained how it could be improved further. This approach can inspire other researchers to use contextual information for classification in other domains and thus define a similar theoretical concept in other domains. We also contributed to customer integration research by demonstrating the value of data derived from customer fleets and demonstrating how geospatial performance measures can be used to develop product and business model innovations by enriching customer usage data.

In practice, we can develop customer-centric and personalized products by using customer data. The use of customer usage data focuses on the pure usage frequency of the entire customer, and less attention is paid to the intention of the customer to use the product. This study shows the high potential of customer usage data in product development. Consequently, more customer-oriented functions can be offered, leading to a competitive advantage. Similar models can be used for other car functions and enhance customer experience with appropriate recommendations.

8. Future research

Future work can involve various topics. The models, such as the random forest, can be further calibrated as predictions with good probabilities with the machine learning classifiers are desirable. In this study, we used the default configurations of the ensemble model. Thus, additional hyperparameter tuning of the best ensemble models can be conducted to improve the results. Another direction can be to develop different models since neural networks have increasingly succeeded in classification tasks in competitive data science.

Another research direction to improve the user experience and recommendations is advanced driver assistance systems with personalization and personalized function recommendations. Currently, driver identifiers are unavailable, but in the future, it may be possible to use driver identifiers to address individual customer needs and provide specific functional information. Vehicle user profiles for driver assistance systems, mobility behavior, and driving dynamics can be used as input features. This allows further personalization of their needs.

Furthermore, the time for which a function should be used to lead to a positive experience for the customer should be determined. In our data, the time (experience label) was determined by experts from an automobile manufacturer and via data analysis. With the help of an expert assessment and a user study, the needed activation duration for a positive function experience for a customer can be better specified, and the use of proactive function recommendations can be further improved.

Proactive function recommendations can be used to create new business models, such as pay-per-use options for functions in the automotive industry. Such models can meet customer needs and recommend the right function at the right time. Proactive function recommendations can also offer functions the customer can test, rent, or buy via digital after-sale services. Additionally, studies can explore which customers use a function frequently and whether they rent or sell this function via a digital after-sale service.

9. References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the association for information systems*, *17*(2), 3.

Bader, R. (2013). *Proactive Recommender Systems in Automotive Scenarios* Technische Universität München].

Bader, R., Siegmund, O., & Woerndl, W. (2011). A study on user acceptance of proactive in-vehicle recommender systems. International Conference on Automotive User Interfaces and Interactive Vehicular Applications,

Bojer, C. S., & Meldgaard, J. P. (2021). Kaggle forecasting competitions: An overlooked learning opportunity. *International Journal of Forecasting*, *37*(2), 587-603.

Bruss, M., & Pfalzgraf, A. (2016). Proactive Assistance Functions for HMIs through Artificial Intelligence. *ATZ worldwide*, *118*(12), 40-43.

Cami, B. R., Hassanpour, H., & Mashayekhi, H. (2019). User preferences modeling using dirichlet process mixture model for a content-based recommender system. *Knowledge-Based Systems*, *163*, 644-655.

Canbek, G., Sagiroglu, S., Temizel, T. T., & Baykal, N. (2017). Binary classification performance measures/metrics: A comprehensive visualized roadmap to gain new insights. 2017 International Conference on Computer Science and Engineering (UBMK),

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* Massachusetts Institute of Technology].

Einizadeh, A., & Kasraei, A. (2021). Proposing a model of green supply chain management based on new product development (NPD) in auto industry.

Enkel, E., Kausch, C., & Gassmann, O. (2005). Managing the risk of customer integration. *European Management Journal*, *23*(2), 203-213.

Guinea, M., Litton, I., Smiroldo, R., Nitsche, I., & Sax, E. (2020). A proactive context-aware recommender system for in-vehicle use. Proceedings of the 2020 4th International Conference on Vision, Image and Signal Processing,

Homola, D., Micus, C., (2023). A computer-implemented method for providing a function recommendation in a vehicle, and a system for implementing the method. https://worldwide.espacenet.com/patent/search?q=pn%3DEP4230491A1

Huang, S., Wang, G., Lei, D., & Yan, Y. (2022). Toward digital validation for rapid product development based on digital twin: a framework. *The International Journal of Advanced Manufacturing Technology*, 1-15.

Jones, M. (2019). What we talk about when we talk about (big) data. *The Journal of Strategic Information Systems*, *28*(1), 3-16.

Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS international transactions on computer science and engineering*, *30*(1), 25-36.

Koushik, S., & Mehl, R. (2015). The automotive industry as a digital business. *Management Summary, NTT Data*.

Lefèvre, S., Carvalho, A., Gao, Y., Tseng, H. E., & Borrelli, F. (2015). Driver models for personalised driving assistance. *Vehicle System Dynamics*, *53*(12), 1705-1720.

Micci-Barreca, D. (2001). A preprocessing scheme for high-cardinality categorical attributes in classification and prediction problems. *ACM SIGKDD explorations newsletter*, *3*(1), 27-32.

Micus, C., Schramm, S., Boehm, M., & Krcmar, H. (2023). Methods to analyze customer usage data in a product decision process: A systematic literature review. *Operations Research Perspectives*, 100277.

Micus, C., Smeets, J., Böhm, M., & Krcmar, H. (2022). Customer Integration in Product Development using Big Data: An Example of Using Fleet Data to determine Parking Behavior.

Nothdurft, F., Ultes, S., & Minker, W. (2015). Finding appropriate interaction strategies for proactive dialogue systems—an open quest. Proceedings of the 2nd European and the 5th Nordic Symposium on Multimodal Communication, August 6-8, 2014, Tartu, Estonia,

NTT-Data. (2020). Connected Car: Consumer Expectations, Opportunities and Challenges for the Industry. Retrieved 12.07.2022, from https://de.nttdata.com/files/2020-en-study-connected-car-expectations-opportunities-and-challenges.pdf

Panniello, U., Tuzhilin, A., & Gorgoglione, M. (2014). Comparing context-aware recommender systems in terms of accuracy and diversity. *User Modeling and User-Adapted Interaction*, *24*(1), 35-65.

Parmenter, D. (2015). *Key performance indicators: developing, implementing, and using winning KPIs*. John Wiley & Sons.

Raschka, S. (2014). An overview of general performance metrics of binary classifier systems. *arXiv preprint arXiv:1410.5330*.

Rokach, L. (2010). *Pattern classification using ensemble methods* (Vol. 75). World Scientific.

Wang, X., Wang, Y., Tao, F., & Liu, A. (2021). New paradigm of data-driven smart customisation through digital twin. *Journal of manufacturing systems*, *58*, 270-280.