Karri Koivula

Discovering the potential of utilizing artificial intelligence in tax procedures

AI-powered artifact as a knowledge creation instrument

School of Technology and Innovations
Master's thesis in Information Systems Science
Master's Programme in Digital Business Development

Vaasa 2022

UNIVERSITY OF VAASA
School of Technology and Innovations

Author:	Karri Koivula
Title of the Thesis:	Discovering the potential of utilizing artificial intelligence in tax procedures: AI-powered artifact as a knowledge creation instrument
Degree:	Master of Science in Economics and Business Administration
Programme:	Digital Business Development
Supervisor:	Ahm Shamsuzzoha
Year:	2022 **Pages:** 83

ABSTRACT:

Artificial intelligence, machine learning, and deep learning have become ubiquitous concepts. Interest in their utilization opportunities in many sectors has exponentially grown during recent decades partly due to the exponential growth of computer power and the increased availability of data, allowing for more powerful and sophisticated information technology solutions. Technological maturity has lowered the threshold, and various open-source libraries and active communities enable the utilization of algorithms such as neural networks in practice. This thesis set out to find whether deep learning algorithms could be utilized in a value-adding way in the procedure for limited liability companies responsible for handling tax claims in the case organization the Finnish Tax Administration. Additionally, the creation and deployment of artificial intelligence solutions should consider legal and ethical manners as restrictive key concerns.

The research was carried out according to the action design research method in which the focus of the research is concurrently building a suitable artifact for the organization and learning (design principles) from the creation and intervention itself. The research method was chosen due to its inclination towards authenticity in the organization and organizational centricity. As a result, the project team consisting of three members created two functional artifacts: one based on neural networks and another based on self-organizing maps. The case organization provided data fueling the deep learning algorithms. Data consisted of financial information of anonymous limited liability companies in Finland. The artifacts were limited to function only as knowledge creation instruments due to legal and ethical limitations present in the context. Knowledge creation in this research context refers to the artifact's ability to identify customers not returning (defaulting) their income tax returns from others.

The created artifacts functioned sufficiently, and their ability to identify defaulting customers from others was promising. Results suggest that it is recommendable to approach problems with more than one artifact solution, and focused roles in the project team are recommended. Artificial intelligence-based artifacts are seen as value-adding since the knowledge created by them can potentially save time, liberate resources and expedite processes. However, finalized artifacts were not created, and testing was limited to a simulated environment. The design principles that emerged from the artifact creation focused on addressing the legal and ethical challenges associated with artificial intelligence in taxation to secure sustainable artifact creation and usage. Design principles were divided into three levels: trustworthiness through accuracy, legal and ethical restrictions and limitations of use, and justification of use. An organization-defined performance threshold needs to be reached by an artifact. An artifact must be transparent and regulated to fulfill context-specified legal and ethical limitations. Lastly, a preliminary inspection of artificial intelligence usage in a case organization is required. Consequently, the preliminary results of this research should be validated by applying the concept in a case organization, followed by an analysis of the results in an end-user setting.

KEYWORDS: Machine learning, neural networks, self-organizing maps, tax procedures

TIIVISTELMÄ:

Tekoäly, koneoppiminen ja syväoppiminen ovat muodostuneet kaikkialla läsnäoleviksi käsitteiksi. Kiinnostus niiden hyödyntämispotentiaaliin monilla toimialoilla on kasvanut viimeisten vuosikymmenten aikana. Laskentatehon ja saatavilla olevan tiedon eksponentiaalinen kasvu mahdollistavat tehokkaampien ja monimutkaisempien ratkaisujen luomisen. Teknologian maturiteetin kypsyminen on laskenut kynnystä ja avoimet ohjelmistokirjastot sekä aktiiviset yhteisöt mahdollistavat neuroverkkojen kaltaisten algoritmien hyödyntämisen käytännössä. Tämän opinnäytetyön tarkoitus oli tutkia tuottaako syväoppimisalgoritmien hyödyntäminen lisäarvoa osakeyhtiöiden verotuksen oikaisumenettelyssä Verohallinnossa. Lainmukaisten ja eettisten tekoälysovellusten luominen ja käyttöönotto tunnistettiin rajoittavaksi ja keskeiseksi tekijäksi.

Tutkimus toteutettiin toiminnan suunnittelututkimuksen mukaisesti, jossa on tarkoitus samanaikaisesti luoda kohdeorganisaation soveltuva artefakti sekä oppia (suunnitteluperiaatteet) artefaktin luomisesta ja interventiosta organisaatioon. Tutkimusmenetelmä valittiin sen organisaatiokeskeisyyden ja organisaatiokohtaisen aitouden vuoksi. Tutkimusmenetelmän soveltamisen seurauksena kolmehenkinen projektiryhmä loi kaksi toimivaa artefaktia, joista toinen pohjautui neuroverkkoihin ja toinen itseohjautuviin karttoihin. Kohdeorganisaatio toimitti syväoppimisalgoritmien tarvitseman datan. Data koostui tunnistamattomien suomalaisten osakeyhtiöiden taloustiedoista. Artefaktit oli rajattu toimimaan ainoastaan nk. tietoa tuottavina työkaluina johtuen lain ja etiikan rajoitteista. Tiedon tuottamisella tutkimuskontekstissa viitataan artefaktin kykyyn tunnistaa asiakkaita, jotka eivät täytä niiden tuloverotuksen veroilmoitusvelvollisuutta.

Luodut artefaktit toimivat riittävällä tasolla. Niiden kyky tunnistaa haluttua asiakasryhmää oli lupaava. Tulosten perusteella on suositeltavaa lähestyä ongelmia luomalla useita erilaisia tekoälysovellutuksia. Lisäksi suositellaan kiinnittämään huomiota keskitettyihin rooleihin projektiryhmässä. Tekoälypohjaiset artefaktit nähdään lisäarvoa tuottavina. Niiden tuottaman tiedon perusteella on mahdollista säästää aikaa, vapauttaa resursseja ja nopeuttaa prosesseja. Viimeisteltyjä ja organisaatioon vapautettuja artefakteja ei luotu. Artefaktien luonnin ja testauksen perusteella syntyneet suunnitteluperiaatteet keskittyivät vastaamaan lain ja eettisyyden asettamiin rajoitteisiin, jotka liittyvät tekoälyn hyödyntämiseen verotuksessa. Näin on mahdollista varmistaa kestävä tapa luoda artefakteja ja ottaa niitä käyttöön. Suunnitteluperiaatteet jaettiin kolmeen tasoon: luottamus tarkkuuden kautta, lain ja eettisyyden luomat rajoitteet käytössä ja tekoälyn käytön perustelu. Artefaktin tulee ylittää organisaatiokohtainen kynnys suorituskyvylle. Artefaktin tulee olla läpinäkyvä ja säännelty, jotta se noudattaa kohdeympäristönsä rajoitteita. Ennakollinen tutkimus tekoälyn hyödyntämiskohteista organisaatiossa on kehoitettavaa. Tämän työn saavuttamat ennakolliset tulokset on suositeltavaa vahvistaa kohdeorganisaatiossa, jota seuraa tulosten analysointi loppukäyttäjien keskuudessa.

AVAINSANAT: Koneoppiminen, neuroverkot, itseorganisoituvat kartat, verotusmenettelyt

Contents

Figures

Tables

Abbreviations

ADR	Action Design Research
AI	Artificial Intelligence
ANN	Artificial neural network
AR	Action research
CM	Confusion matrix
DL	Deep learning
DR	Design research
FN	False negative
FP	False positive
KBS	Knowledge-based system
LLC	Limited liability company
LR	Linear regression
ML	Machine learning
MLP	Multi-layer perceptron
RL	Reinforcement learning
ROC	Receiver Operating Characteristic
SL	Supervised learning
SOM	Self-organizing map
TN	True negative
TP	True positive
UL	Unsupervised learning

1 Introduction

The first chapter presents the background of the study, including the motivation behind it and a description of the problem domain. The chapter introduces the research problem, goal, objective, and questions. The chapter proceeds then to the study's limitations and ends with the structure of the study.

1.1. Background and purpose of research

The Finnish Tax Administration's main tasks include but are not limited to carrying out taxation and related payments, tax control, and recovery of unpaid taxes. Tax administration is accountable to the ministry of finance. (Tax administration act (2010/502) 1 § and 2 §) Tax control and how it should be done are not described explicitly in the law. Tax control's main objectives include reducing the tax gap, combating the shadow economy preventatively and precisely, and collaborating with other authorities. Tax control is issued to everyone from the individual to the corporate level. (Finnish Tax Administration, 2019a)

Long processing times in the claim for adjustment procedure in the tax administration were identified as a research opportunity. Claim for adjustment procedure refers to the processing of tax claims (Tax assessment act, chapter 5). This research focuses on a specific group of customers within the claim for adjustment procedure: limited liability companies.

Due to long processing times, claims for adjustment is viewed as a bottleneck procedure. Partially responsible for the long processing time is the amount of unnecessary cases created by corporate taxpayers not filling their tax obligations within a set time limit. By not returning a tax return, a tax year is estimated. Estimating a tax year means taxes are enforced based on an estimate and not on the company's actual tax report. Reversing

the estimated tax years is time-consuming for the tax administration. The potential identification of such examples would be beneficial as knowledge to help pre-emptively tackle similar cases in the future. Decreasing the number of estimated tax years is helpful for customers and the tax administration. This research aimed to create a knowledge-creating artifact based on AI that could help the Finnish Tax Administration decrease the number of estimated tax years pre-emptively and concurrently make a stronger sense of the potential usage possibility of such artifacts. Vast amounts of data available in the tax administration create a suitable environment for AI solutions to be developed. Information created by AI solutions can decrease the amount of manual work. Additionally, as the claim for adjustment procedure strictly follows laws and regulations, the artifact and its creation focus on recognizing examples that go through the procedure and not participating in the decision-making of any sort.

Advancements in computer power, the exponential growth of available data, and the undeniable potential within artificial intelligence (AI) facilitated the idea for this thesis. AI solutions based on artificial neural networks (ANN) and self-organizing maps (SOM) were used to extract valuable information related to a bottleneck process in the tax administration. (Finnish Tax Administration, 2021a) This thesis aims to prove that AI could bring value to taxation and can extract useful information from the data.

Applying for a claim for individual and corporate customers is a fundamental right. However, many of the claims are tax returns that have not been returned in the first place. By filing the tax return significantly late (10 months after the accounting period has ended), the processing time is approximately 12 months due to several reasons, including manual processing, which is not optimal for the taxpayer or the tax administration.

Two different AI solutions, IT artifacts, were built according to the action design research (ADR) research method. The artifacts were created to function as pre-emptive knowledge-creating instruments. Additionally, the ADR research method generated design principles to make such artifacts in similar settings.

This thesis is interested in finding out if an AI solution could detect and recognize the underlying characteristics in corporate taxpayers that would explain the number of late filings. Reducing the number of late filings would benefit the taxpayers as they would receive their tax decisions more swiftly. The tax administration would improve the quality and reduce the number of claims. ANN and SOM are applied to the available data to see how well it can differentiate a company that will file its tax return significantly late from one that does not.

This thesis answers whether AI could add value to taxation and what challenges might be faced when utilizing AI solutions in taxation. Answers to these questions are provided based on the results achieved by the ADR project carried out in the case organization.

1.2. Research problem and goal

The research problem is whether AI could add value to taxation in the given problem domain. The research goal is to answer the research problem of how AI could be applicable in taxation in the limited liability companies' (LLC) claim for adjustment procedure. The research objective states that a suitable IT artifact would be created with the case organization to address the research problem and reach the research goal. Performance and creation are afterward analyzed.

The creation of the artifact follows the ADR research method. The researcher formed research questions before the creation of the IT artifact began. Research questions are defined as:

1. How can AI be deployed to the case organization in Finland to create value in its current taxation system?

The first research question revolves around what kind of value and information AI creates in the case organization and how it can be achieved. Reasons and timing of the usage are considered and planned accordingly.

2. How can AI be deployed so that it does not violate rules and regulations?

The second research question focuses on challenges that AI usage should consider, such as trustworthiness, legality, and ethical restrictions, and how they are addressed.

1.3. Limitations

This research only focuses on utilizing NN and SOM as a pre-emptive data analysis tool in the Finnish Tax Administration. In the context of this research, artificial neural networks (ANN) and SOM are part of deep learning (DL). Data used in this research is limited only to limited liability companies with a business source of income. The companies' data is mainly taxation and accounting-related basic and business information.

1.4. Research structure

The introduction chapter presents the research problem, limitations, structure of the thesis, and the core of the idea to the reader. The second chapter focuses on machine learning (ML) and DL, which form the primary theoretical groundwork for this thesis. In the third chapter, the reader is introduced to LLCs and their taxation procedure in general. The fourth chapter presents the research method, and the fifth chapter outlines the available data used in this thesis. In the sixth chapter, the reader learns how the utilization of ADR was implemented, how the artifacts performed, and what were the design principles that came out of the process. Moreover, the sixth chapter provides answers to the research questions. The seventh and final chapter discusses the research's goal

and objectives, covering how this research study was evaluated and the related restrictions. The final chapter also includes necessary suggestions for future research.

2 Theoretical framework

This chapter consists of the main theoretical background related to AI and its main concepts of ML and DL. This chapter ends by presenting performance evaluation metrics used in assessing the performance of IT artifacts created in this study.

2.1. Artificial intelligence

Akerkar (2019, p. 3-4) views AI as the replication of biological, analytical, and decision-making capabilities as the essence of artificial intelligence. AI is often defined *as* "the science and engineering of imitating, extending and augmenting human intelligence through artificial means and techniques to make intelligent machines." To be considered intelligent, a system should be able to learn in a changing environment (Alpaydin, 2014, p. 3).

The "Dartmouth conference" in 1956 is considered the official start of AI as it marked the beginning of AI as a research field (Shi, 2011, p. 2). To some extent, the famous Turing test by Alan Turing even precedes the "Darthmouth conference" by offering a view on how to identify an "intelligent machine" (Rahman, 2020, p. 15-17; Shi, 2011, p. 2). Artificial intelligence (AI) is divided into general and narrow categories. General AI refers to AI that can act in a "human intelligent way," navigating different problem domains. The ability to adapt to ever-changing situations is referred to as general AI. Currently, there are no systems that can perform this way. Narrow AI is an application that can perform well in one or two things but cannot go beyond what it has not been designed for, such as an AI application used to detect tax avoidance. The AI of today is in the narrow category. (Akerkar, 2019, p. 3-4)

Finlay (2018, p. 62) offers a general way of viewing AI problems by dividing them into two broad categories: simple and complex. Simple problems have a singular objective

that must be determined and is easier to quantify. According to Alpaydin (2014, p. 5), a classification task with two separating classes is an example of such a problem. Whether a person is admitted a bank loan or not is an example of a classification problem. Finlay (2018, p. 62) highlights that a complex problem has more than one objective. Problems that require multiple ML approaches combined, such as autonomous vehicles, are complex AI problems.

Knowledge-based systems (KBS) are perhaps AI's most successful practical branch (Akerkar, 2019, p. 4). According to Shi (2011, p. 25), "KBS includes expert systems, knowledge base systems, intelligent decision support systems.." KBS systems primarily consist of a knowledge base and an inference engine. The knowledge base includes facts, task-related specifics, and heuristic knowledge of the domain. Inference engine refers to various methods of deducing new information from the knowledge base (Shi, 2011, p. 25, 120; Benfer et al., 1991, p. 11).

2.1.1. Machine learning

"ML is the systematic study of algorithms and systems that improve their knowledge or performance with experience" (Flach, 2012, p. 3). According to Finlay (2018, p. 12), the main ingredients that fuel most AI and ML applications include data input, data preprocessing, predictive models, decision rules, and output. Due to exponential growth in computational power and the availability of a vast amount of data (big data), learnings methods such as ML and DL have become more attractive (Alpaydin, 2014, p. 309). In ML, we are interested in discovering patterns and useful approximations from data (Alpaydin, 2014, p. 2). Data input can be almost anything from sensory inputs such as videos to filed online forms such as tax returns. Data preprocessing refers to turning data inputs into a computer-friendly format. (Finlay, 2018, p. 12)

Jung (2018, p. 9) divides ML into three components: (1) Data, its features, and labels, (2) a model or hypothesis space, and (3) loss function. Data is viewed as a collection of data

points that contain information of any kind. The amount and quality of data are crucial for ML. Features are measurable properties of data (Mirjalili & Raschka, 2019, p. 9, 109). An essential part of ML is to figure out features that have the most significant effect on the performance of ML. Data is often but not always labeled. Labels refer to higher-level information, and like features, they characterize a data point. ML's model (hypothesis space) is a restricted, computationally feasible map of label and feature space. The map is called either a predictor or classifier. The map is called a classifier for finite label spaces, and for continuous label space, the map is called a predictor. (Jung, 2018, p. 7)

According to Jung (2018, p. 9), linear regression is (LR) a supervised machine learning method that uses linear maps for the hypothesis space. LR tries to find a map that could predict an accurate label of an output based on features of a data point. To acquire such a map, historical data is used to try out different options for the map and pick the best one. The purpose of the loss function is to measure the quality of a specific map. "Loss (approximation error) is the sum of losses over the individual instances." (Alpaydin, 2014, p. 41-42). To determine the feasibility of the map, a measurement for the loss (or error) incurred needs to be specified. For the LR example involving numeric labels (regression problem), a commonly used choice for loss function is the squared error (R^2) loss. (Jung, 2018, p. 26)

ML requires data as its primary goal is to predict an outcome based on features (Giussani, 2020, p. 14). Received data almost always needs to be prepared before being used in ML (Lee, 2019. p. 107-117). A case-dependent number of iterations of training, validating, and testing before a finished ML algorithm can be utilized in action, which includes splitting the dataset into testing and training sets, data feature-related selections, and dimensionality reduction (Chebbi, 2018, p. 213; Campesato, 2020, p. 28-34).

Knowledge representation refers to the fundamental goal of AI; the creation of such AI that is capable of intelligent behavior as determined by humans (Shi, 2011, p. 18). Machine learning is a "facet of AI that focuses on algorithms, allowing machines to learn

without being programmed and change when exposed to new data." (Akerkar, 2019, p. 4). ML is seen as the most critical problem of AI (Shi, 2011, p. 18). According to Taiwo (2010, p. 4-5), ML is suitable for tasks that cannot be defined well except by example. ML can roughly be divided into supervised, unsupervised, and reinforcement learning. (Alpaydin, 2014, p. 9-13)

2.1.1.1.   Supervised learning

According to Campesato (2020, p. 19), in supervised learning (SL), learning is done by exposing the learner to the data, including the known outcomes. This way, the machine can improve its performance; thus, it knows the desired results and what to pursue. Learning in SL occurs during training. (Rahman, 2020, p. 20) According to Alpaydin (2014, p. 5), both regression and classification are supervised learning problems.

2.1.1.2.   Unsupervised learning

Unsupervised learning (UL) includes only the input as output data is missing or excluded (Alpaydin, 2014, p.11). Unsupervised learning includes unlabeled or data of unknown structure. It is used to deduce important information from data to find patterns. (Lee, 2019, p. 5) Unsupervised learning is suitable for finding regularities in the data and detecting naturally occurring groups, such as in the k-means clustering algorithm (Alpaydin, 2014, 11; Rahman, 2020, p. 21).

Additionally, the self-organizing map (SOM) is a well-known UL method. SOM draws a topographical map of the data where similar observations are positioned closer, and an ordered representation of the data is created. As Kohonen (2013, p. 52-53) presented, figure 1 depicts the core idea of SOM: input data mapped out where Mc best represents X. Models (Mi) in the same circle are more similar to Mc than M to other observations on the map.
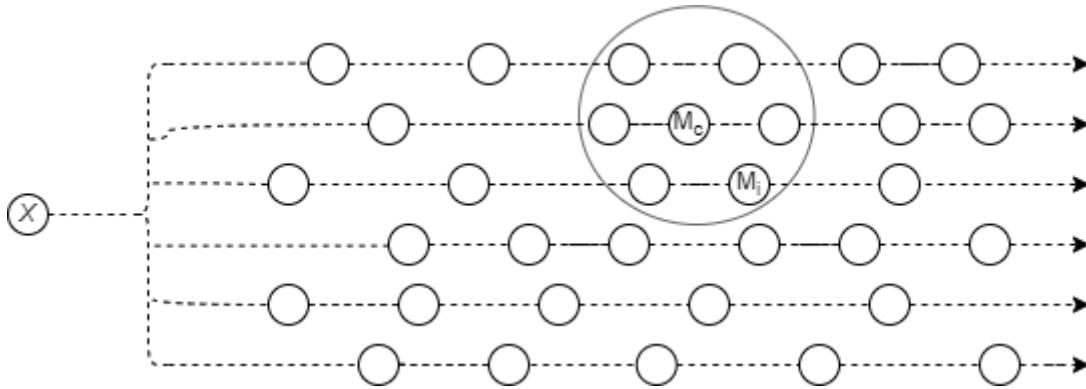
Figure 1. Illustration of SOM

2.1.1.3.   Reinforcement learning

According to Alpaydin (2014, p. 517), in reinforcement learning (RL), a "decision-making agent" is acting in an environment from which it receives feedback (reward or penalty) when trying to solve a task. Based on the feedback, the agent should be able to learn the best policy for acting in the environment. Learning the policy is at the center of RL. Individual action is determined well if it supports the longer-term goal, such as a chess move to win the game (Rahman 2020, p. 80).

RL is commonly utilized in games due to its nature as a series of sequences or actions towards a goal. More than one agent is possible in tasks where concurrent action is required. In cases of multiple agents, the agents communicate and cooperate to complete a task. (Dutta, 2018, p. 47-48)
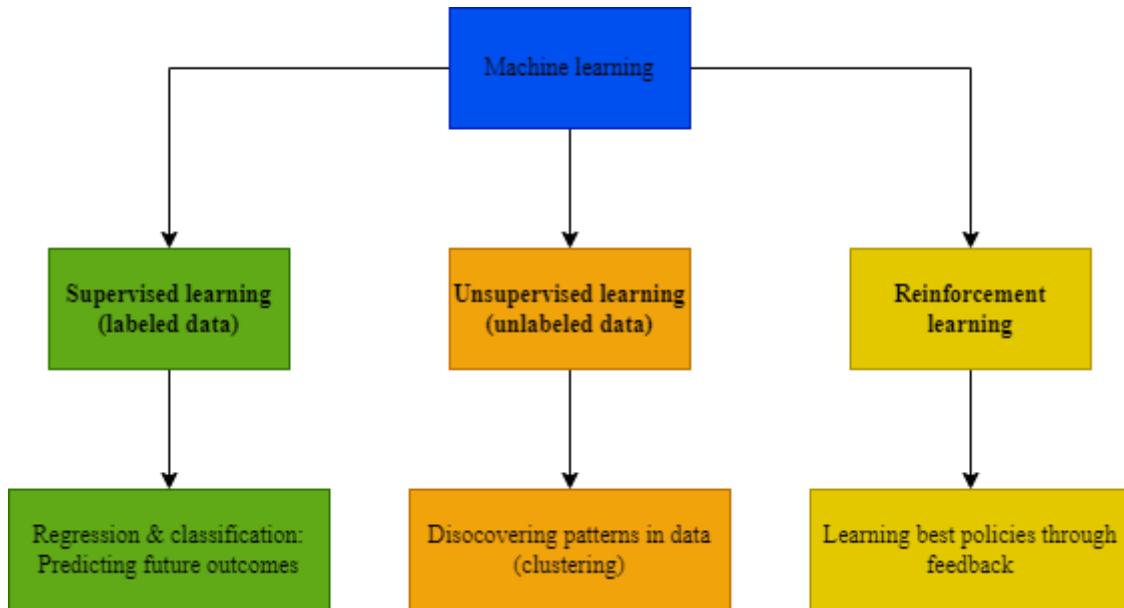
Figure 2. SL vs. UL. vs. RL

Figure 2 depicts the three main learning categories in ML and their central learning policies.

2.1.2. Deep learning

DL is viewed as a subfield of ML that utilizes multiple layered ANNs to solve problems (Mirjalili & Raschka, 2019, p. 383-384). Instead of analyzing data linearly, neural networks enable machines to process data nonlinearly (Alpaydin, 2014, p. 306). At its core, DL divides the learning process into connected steps, also known as layers, that are assigned to different sections of the main problem available to the whole network (Rahman, 2020, p. 80-81). According to Kelleher and Tierney (2018, p. 242), the strength of DL models lies in their ability to utilize previously gathered knowledge from the previous layers to their advantage in the following layers, which is referred to as backpropagation. In backpropagation, previously accumulated feedback from events in the network is used in future calculations within the network (Rahman, 2020, p. 22).

2.1.2.1. Artificial neural networks

An ANN mimics the human brain and its functions; hence the neural in the neural network refers to biological neurons in the brain (Graupe, 2013, p. 1). A neural network consists of layers that contain neurons that perform the required mathematical calculations (Rahman, 2020, p. 20). The neurons in the layers together form a parallel and interconnected network as each of the layers and their neurons might connect (Rahman, 2020, p. 20; Alpaydin, 2014, p. 267). UL algorithm SOM is also considered a type of ANN (Kohonen, 2016, p. 724).

According to Mirjalili & Raschka (2019, p. 83), the three prominent neural networks models include (1) feedforward neural network (FNN), (2) recurrent neural network (RNN), and (3) convolutional neural network (CNN). In FNN, the connections in the network are only moving forward (Kelleher & Tierney 2018, p. 124). In RNN, the neurons are also connected backward, resulting in the network having a short-term memory of past incidents (Mirjalili & Raschka, 2019, p. 83-84; Alpaydin, 2014, p. 305). RNN is utilized when the network is required to know the information of the previous layers. In CNN, "the work of each hidden unit is considered to be a convolution of its input" (Alpaydin, 2014, p. 294). Hidden units in CNN view the same input space from a different place, looking for additional features that are later intertwined into more useful information. CNN is used in visual recognition tasks. (Alpaydin, 2014, p. 295; Rahman, 2020, p. 84)

2.1.2.2. Multilayer perceptron

The multilayer perceptron is a feedforward neural network model as all the connections move towards the output (Kelleher & Tierney 2018, p. 124). Perceptron refers to the ensemble of a neuron and its input connections and weights (Alpaydin, 2014, 271, 273). There are some constant parameters to consider when training a multilayer perceptron network, such as the number of hidden layers in the network as the increased number of hidden layers makes the network "deep", activation function, or the calculation of a

neuron's activation threshold, and the batch size or size of the data section is passed to the network in the training phase (Jung, 2018, p. 45). Additionally, epochs or the number of passings of the data through the network (Alpaydin, 2014, p. 285) and learning rate or how quickly the network optimizes itself (Mirjalili & Raschka, 2019, p. 200).
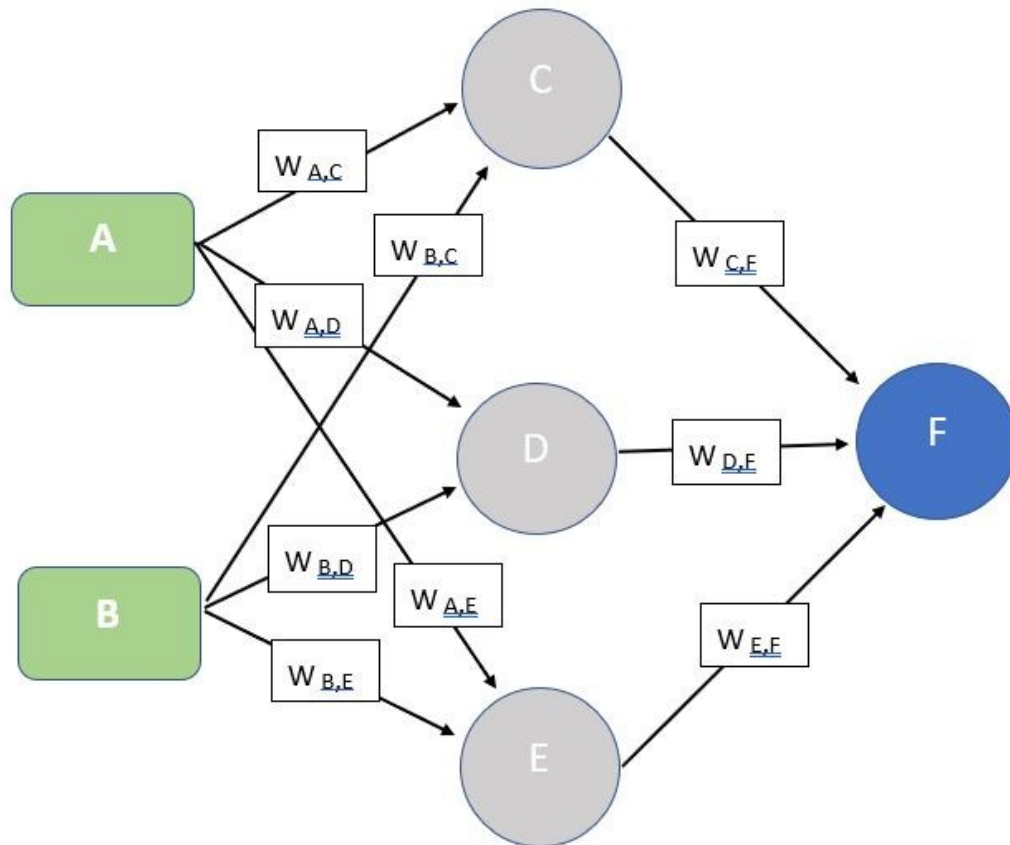
Figure 3. Feedforward neural network

In figure 3 (adapted from Kelleher & Tierney 2018, p. 124), there are three layers of neurons: (1) input layer, A and B, (2) hidden layer C, D, and E, and (3) output layer F.

Neurons in a neural network are doing a set of operations:
1. Multiplying each input by a weight
2. Adding together the results of the multiplications
3. Pushing the result through an activation function

According to Kelleher and Tierney (2018, p. 121-136), "all the connections between the neurons in a neural network are directed and have a weight associated with them." The weight applied to an input that a neuron receives is the weight on the connection coming to the neuron when the multi-input regression function over its inputs is calculated. As seen in figure 3, the flow of information in the network between the neurons is presented by arrows. The neural network in figure 3 is considered fully connected because each neuron is connected to all the neurons in the subsequent layer. The tags in the arrows reveal the weight that the neuron at the end of the arrow applies to the information passing through the connection. In figure 3, the calculation performed by neuron F of the network can be defined as:

$$Output = \varphi(\omega_{C,F}C + \omega_{D,F}D + \omega_{E,F}E)$$

* φ = activation function

** ω = weight applied to the neuron

2.1.2.3.   Predicting future events in taxation using DL models

Tax officials in Finland have already utilized and expressed a growing interest in AI usage in taxation (Finnish Tax Administration, 2021b). Chen et al. (2011) developed an automatic detection model for discovering erroneous tax reports in their study. The study was motivated by the criticality of tax reporting and the large number of errors found in reports in recent years. Detecting erroneous tax reports is tedious and depends on experienced personnel. Therefore, the need for an automatic solution exists to reduce the workforce needed for the job. The model in the study by Chen et al. (2011) was carried out with various NN methods compared to each other. The different approaches were "multi-layer perceptrons, learning vector quantization, decision tree, and hyper-rectangular composite neural networks methods." Data consisted of construction companies residing in Taiwan. No matter which NN approach was used, the correct recognition rate reached nearly 80 %. The best performing approach, hyper-rectangular

composite neural network, was able to digest almost 250 valuable rules for identifying erroneous tax reports from the data.

Studies by Xiangyu et al. (2018) and Pérez López et al. (2019) focused on tax evasion. Xiangyu et al. (2018) developed a neural network model to tackle the issue of tax evasion in automobile sales enterprises in China. The NN-based recognition model's object was to determine behavior related to tax evasion. Pérez López et al. (2019) utilized in their research an MLP neural network model to identify tax fraud concerning personal income tax returns in Spain. The result in both cases was a success. Xiangyu et al. (2018) reached a recognition accuracy of 89 %. The result was assessed with Receiver Operating Characteristic (ROC) curve, which showed that the classification effect was good. Pérez López et al. (2019) achieved an efficiency rate of 84.3%.

Moreover, the NN by Pérez López et al. (2019) offered information on the probability of each taxpayer's inclination to evade taxes. MLP is beneficial for classifying fraudulent/non-fraudulent taxpayers based on the results. The robustness of the model was confirmed with the ROC curve, which verified the NN's high predictive capacity.

A study by Rahimikia et al. (2017) focused on tax evasion with a more complex approach. In their study, Rahimikia et al. (2017) created a novel hybrid intelligent system to detect corporate tax evasion in Iran. Hybridity came from combining NN, SVM, and LR classification models with harmony search (HS) optimization algorithm, which is inspired by the improvisation process of musicians. The system was tested in the food and textile sectors. Researchers concluded that the system could accurately detect hidden patterns in tax returns that could point toward tax evasion. The results offer valuable, sector-wise information about the financial structure of tax evasion. The hybrid system is seen as a useful tool to detect tax evaders and an identifier of patterns suggesting tax evasion.

Additionally, tax officials have utilized NNs in social media. Zhang et al. (2020) developed a proof-of-concept NN to identify transaction-based tax-evading activities in the hidden

economy of social media. Dataset consisted of "Instagram posts about #lipstick and manually annotated sampled posts with multiple labels related to sales and tax evasion activities." The purpose of the NN detection model was to identify suspicious social media posts. The posts deemed more suspicious by NN were afterward analyzed by tax officials. As the NN model identifies the suspicious posts, first, the productivity of manual work is improved from 22 percent to 72 percent. The NN model improves manual labor efficiency as the tax officers will not have to select the posts randomly.

2.1.3. Performance evaluation metrics

Several ways exist to measure the performance of a neural network model. The performance of the neural network created in this thesis is evaluated with the help of accuracy, precision, recall, and f1-score, all derived from the confusion matrix.

The confusion matrix (CM) presents the performance of a learning algorithm. CM is a square matrix that reports the count of true positive (TP), true negative (TN), false positive (FP), and false negative (FN), as presented in figure 4 (adapted from Rokach (2009, p. 160-161)). Table 1 illustrates how they are calculated. (Adapted from Giussani, 2020, p. 62-64).

Figure 4. Confusion matrix

Table 1. Evaluation metrics

Metric	Calculation	Definition
Accuracy	$\dfrac{TP + TN}{FP + FN + TP + TN}$	Sum of correct predictions divided by all predictions.
Precision	$\dfrac{TP}{TP + FP}$	Correctness of the model per class. Useful in imbalanced class problems.
Recall	$\dfrac{TP}{FN + TP}$	The number of correctly evaluated instances per class is divided by all the correct examples in the class. Useful in imbalanced class problems.
F1-score	$\dfrac{2 \times Precision \times Recall}{Precision + Recall}$	A balanced combination of precision and recall.

3 Limited liability company and its taxation procedure

This chapter maps out an overview of what an LLC is and what is expected from them in terms of accounting and taxation and an overview of the taxation procedure of LLCs in Finland, focusing on automation, monitoring, and ethical principles for AI in the Finnish Tax Administration. Legislative changes concerning LLCs that came into force in 2020 or after are not considered as the data used in this study is from tax years 2017, 2018, and 2019.

3.1. Principles of LLC's taxation

According to the limited liability companies act (2006/624), chapter 1, 1 §, subsection 1, LLC is a separate taxpayer from its owners created through registration to the trade register. By registering to the Finnish trade register LLC becomes a discrete taxpayer. Stockholders are not personally responsible for LLC's liabilities (Limited liability companies act chapter 1, 2 §, subsection 2).

According to the act on bookkeeping (1997/1336) chapter 1, 1 § subsection 1 paragraph 1 LLCs are accounting obligated for each accounting period. LLCs' are obliged to compose a financial statement, including a balance sheet, profit/loss (P/L) statement, and annual report for each accounting period, according to the act on bookkeeping and limited liability companies act (Limited liability companies act chapter 8, 3 §). The accounting period is 12 months except for the beginning and end of the business and in cases of alterations to the accounting period (Act on bookkeeping chapter 1, 1 §, subsection 1). The balance sheet dictates the corporation's financial status: the relationship between assets and liabilities. P/L statement presents how the accounting period's outcome came to be. In some cases, LLCs' are required to submit separate funds statement which dictates how funds were acquired and utilized (Act on bookkeeping chapter 3, 1 §, subsection 1, paragraph 1, 2, and 3).

Larger corporations such as public LLCs are expected to return an annual report which is a written report of a company's status on development and profitability, financial situation, and most significant risks (Act on bookkeeping chapter 3, 1 §, subsection 3). A financial statement is required to present a genuine, essential, and sufficient picture of the profitability and economic status of the company (Act on bookkeeping chapter 3, 1 §, subsection 1). Corporations' tax obligations are based on the financial statement created according to the act on bookkeeping (Act on bookkeeping chapter 1, 1 a §, subsection 2).

According to the income tax act (1992/1535) section 1, subsection 4, corporations in Finland are tax liable. Corporations consist of governments, municipalities, congregations, limited liability companies (LLC), and foreign estates (Income tax act 3 §, subsection 1). Monetary benefits received by the tax liable are taxable income. The tax liable has the right to deduct expenses related to the acquirement and retaining of the benefits. Therefore, profit or loss is determined by subtracting tax-deductible costs from taxable income (Income tax act 29 §, subsection 1). Corporations' taxable income is calculated separately for each source of income (Income tax act 30 §, subsection 4).

Before 2020, three different income sources exist for LLCs: personal, business, and agriculture. Personal source income is taxed according to the income tax act. Taxation of business source income is done according to the act on the taxation of business income (1968/360). Agricultural source income is carried out according to the act on agricultural income tax (1967/543).

The tax rate percentage for corporations is 20 (Income tax act 124 §, subsection 2). Confirmed losses are deducted in the order they have occurred (Income tax act 117 §, subsection 2). Losses from business activities are deductible from taxable income for ten subsequent years. Losses belonging to the same source of income from the previous tax

years are deducted from the current fiscal year's taxable profit in the same source of income (Income tax act 119 §, subsection 1).

According to the act on assessment of assets in taxation (2005/1142) 2 §, subsection 1, net worth (positive/negative) for a non-public LLC is calculated by subtracting the total amount of liabilities from the total amount of assets. LLC's assets consist of fixed assets and other non-current investments, current assets, financial assets, and other assets with monetary value. Liabilities include borrowed capital in the balance sheet. (Act on assessment of assets in taxation 2 §, sub-section 2 and 3)

The assets in 2017, 2018, and 2019 tax returns consisted of the following fixed assets and other non-current investments, current assets, financial assets, and other long-term investments. The liabilities consist of current and non-current liabilities. Net worth is calculated by subtracting liabilities from assets as presented in adapted table 2. Capital, equity, and reserves are presented as recorded in accounting. (Finnish Tax Administration, 2021c)

Table 2. Compressed adaption of net worth calculation

1 ASSETS	2 LIABILITIES
Fixed assets and other non-current investment	Current liabilities
Current assets	Non-current liabilities
Financial assets	
Other long-term investments (Income Tax Act)	
ASSETS TOTAL	LIABILITIES TOTAL

NET WORTH - POSITIVE
NET WORTH - NEGATIVE

3 CAPITAL, EQUITY, AND RESERVES
Restricted equity
Unrestricted equity
CAPITAL, EQUITY, AND RESERVES TOTAL

3.2. Tax assessment procedure for LLC

Act on tax assessment (1995/1558) is a 12 chapter and 96 section law on taxation procedures and claims for adjustment on income taxation. It dictates taxpayers reporting responsibilities and procedures of tax assessment. Principles of tax assessment procedures are mainly based on the fourth paragraph of the tax assessment act and claims for adjustment in the fifth paragraph.

The Finnish Tax Administration carries out taxation in Finland. Taxation will be carried out based on taxpayers' reporting and reports received from external parties. (Tax assessment act 6 § and 26 §)

According to the act on administrative procedures (2003/434) 1 § purpose of the law is to execute and advance good administration and due process in administrative procedures. The purpose of the law is also to advance the quality and productivity of administrative services, such as the tax assessment procedure and carrying out taxation.

LLCs' are obligated to give their reports (tax return) four months after their accounting period has ended. An LLC which is neglecting its obligation results in taxation being estimated by the tax administration. Tax administration must send a hearing to the taxpayer of the estimation to do this. Taxation will be estimated if the reporting obligation is not fulfilled within the time reserved in the hearing. (Tax assessment act 7 §, 8 §, and 27 §; Tax administration's decision on reporting duties and notes (A123/200/2016))

For corporations, taxation ends at the latest ten months after the end of their tax year closing month (= end of the accounting period). If the taxpayer has not filed their tax return within ten months, the tax decision will be based on the estimate made by the tax administration. To adjust a closed tax year, an adjustment claim is required. Processing time for a closed tax year is 12 months. (Tax assessment act 49 § and 61 §; Finnish Tax Administration, 2021a) Figure 5 demonstrates how the taxation procedure occurs for an LLC.

LLC's accounting period has ended on 31.12.2018 therefore its tax year is 2018.

31.12.2018

-Closing of the accounts

-Tax return filed based on accounting

30.4.2019

Tax return is expected to be returned if not otherwise stated.

Tax assessment

-Tax control by tax administration
-Hearing taxpayer on tax estimation (letter)

31.10.2019

Last day to file tax return and remove estimation

Claim for adjustment

-Only way to remove the estimation is by submitting a claim

-Processing time from 12 to 18 months

31.12.2021

Right to apply for a claim is three years

Figure 5. Tax assessment procedure simplified

The principles of the tax assessment procedure are a guide based on the tax assessment act and administrative procedures. It maps out the basic principles of how taxation is carried out in Finland. The principles of taxation procedures apply to all taxpayers in general. Taxation procedures are primarily based on the tax assessment act. However, administrative procedures can supplement the Tax assessment act if not otherwise stated. (Finnish Tax Administration, 2015)

3.2.1. Automated taxation in Finland

Tax administration is responsible for the tax assessment procedure. As taxation is mainly based on the reporting by taxpayers, it is left to the tax administration to ensure that

taxation is carried out correctly. Reports subject to manual control are selected based on specific rules. (Finnish Tax Administration, 2020)

The amount of tax reports is vast. With over 15 million tax-related decisions, not everything is manually revised. Automated decision-making is necessary due to the immense amount of tax-related work. Many tax-related procedures are carried out automatically without revising by a tax official. (Finnish Tax Administration, 2020)

Automation is directed to undisputed matters which are not selected to manual control and could be solved without consideration. Cases not selected for manual control are formal. Tax decisions made by automation are not explicit and binding decisions made by the tax officials. If an error occurs, it can be corrected afterward by the taxpayer and the tax administration. Tax administration is not utilizing artificial intelligence in tasks that require consideration and decision-making. That is left solely for tax officials. (Finnish Tax Administration, 2020)

All the assessments and procedures made by the tax administration are based on law. The automated decision-making is used when it is possible to program a set of rules based on legislation. The algorithms or logic behind automation have not been strictly defined in legislation. However, automated solutions can only be used in situations that have been mentioned explicitly in the law or procedures based on law. AI, statistics, or scientific models are not used in automation. (Finnish tax administration, 2020)

Safeguard measures to ensure that taxation does not violate the fundamental rights of taxpayers are taken. Measures strictly and only directed to automated taxation do not exist. However, the safeguard measures, in general, apply also to automated taxations. Safeguard measures are the following:

1. Taxpayers are heard before decision-making that requires consideration.
2. Taxpayers have the right to submit a claim that a tax official processes.

3. Right to know when taxation has been assessed automatically.

4. Right to know what is automatically assessed.

5. Right to expect that taxation is processed by a tax official when additional reporting is submitted.

(Finnish Tax Administration, 2020)

Automation has been used in taxation in Finland since 2005 since it has increased the performance of tax-related processing. In 2019 deputy ombudsman of the Finnish Parliament released a decision relating to automation in taxation and its relationship with taxpayers' due process, good administration, and tax officials' liabilities due to several complaints by taxpayers who had problems with automated taxation. (Finnish Tax Administration, 2020; Finnish Parliament Ombudsman, 2019)

According to the deputy ombudsman, automated taxation is not based on appropriate and precise legislation which considers good administration, due process, and the tax official's liabilities; it is therefore against the fundamental law. Immediate investigation of regulative needs is expected. (Finnish Parliament Ombudsman, 2019)

The deputy ombudsman pointed out that legislation behind automated taxation applies only to tax assessment, not tax-related decision-making. Taxation is mainly automated, and many phases from hearing to decision-making, occur in automation without any tax officials taking part in the process. Additionally, the automated process and the algorithms behind it are not transparent. For this to actualize, precisely defined legislation is required. Taxpayers should be openly notified when their taxation has been automatically assessed and how, so they can evaluate if it has been done correctly and their due process is respected. (Finnish Parliament Ombudsman, 2019)

3.2.2. Overview of monitoring in tax assessment

The strategic goals of the Finnish Tax Administration are securing tax funding, righteously carrying out taxation and positive customer service. The automation rate in taxation has increased significantly through the years. New technologies have allowed citizens to fulfill their responsibilities and for the tax administration to validate, assess, and control the flow of reports, payments, and information more efficiently (Finnish Tax Administration, 2019a)

The more routine tasks such as copying information and comparing data have already been partly given to robots. Robots in this context refer to robotic process automation (RPA), suitable for routine tasks and are based on explicit rules in digital form. They have the characteristics of assembly-line work. (Finnish Tax Administration, 2019a)

3.2.3. Ethical principles for AI in the Finnish Tax Administration

In 2018 the Finnish Tax Administration released its ethical principles for AI. The Finnish Tax Administration joins in the pursuit of attaining ethical and trustworthy AI. According to the Finnish Tax Administration, ethical principles will be considered in all decision-making considering AI. (Finnish Tax Administration, 2019c)

The principles of the tax assessment procedure are based on two laws: the legislation on tax assessment and administrative procedures. The ethical principles for AI are based on principles of the tax assessment procedure. The ethical principles of AI consist of four main principles, as presented in table 3.

Table 3. Ethical principles for AI

P1	Reliable data.	How AI solutions function is known, and detailed knowledge of their operating principles has been acquired. AI will not be given access to data until it is certain that the data is reliable and suitable for its purpose. Reliability and suitability will be actively monitored even when data is used. Inaccurate data and algorithms will undergo necessary corrections promptly. (Finnish Tax Administration, 2019c)
P2	A human is always responsible.	AI can be taught by a human, or it can learn by itself though it is continuously monitored by a human. Suggestions created by AI can always be changed. AI can only proceed in the decision process if it can be traced and justified afterward. A facet in charge of AI has been named. (Finnish Tax Administration, 2019c)
P3	AI follows laws and regulations.	Usage is monitored and evaluated. Immediate action will be taken in case of divergence. AI will not endanger taxpayers' tax data security or confidentiality. The use of AI doesn't jeopardize the legal protection of the taxpayer or the person responsible for AI's decision. AI partners are selected with care, and tax administration takes full responsibility for the operations of the entire supply chain. AI solutions undergo the same safety principles as all other IT systems utilized in the Finnish Tax Administration. (Finnish Tax Administration, 2019c)
P4	Tax administration takes part in public discussion on responsible and ethical AI applications.	Adoption of ethically sustainable AI technologies and international procedures is promoted. The tax administration also influences changes in legislation. Tasks in which AI is utilized will be openly communicated to the public. (Finnish Tax Administration, 2019c)

4 Methodology

This chapter consists of the reasoning behind the selection of the ADR research method, the theoretical framework of ADR, and the data collected for the study.

4.1. Justification of methodology

The researcher chose ADR by Sein et al. (2011) as a research method for the project due to its flexibility, authenticity, and organization centricity. ADR team in the project consisted of three professionals working on different tasks within the case organization, including the researcher. Several meetings and altogether three development cycles occurred during the project to create a suitable IT artifact that could assist the procedure by offering new insights.

The problem could be classified in the simple category described by Finlay (2018, p. 62) and Alpaydin (2014, p. 5). ADR outcomes consist of an IT artifact for the organization and design principles through generalized outcomes for the scientific community (Sein et al., 2011, p. 44). As a result of utilizing ADR in the case organization's restricted and authentic setting, the project concluded by creating two different IT artifact solution concepts: NN as an SL type of solution and SOM-based algorithm as a UL type of solution. Moreover, preliminary design principles emerged from the ADR process.

4.2. Action design research

This study adopted action design research (ADR), which is a research method with its research process focused on building innovative IT artifacts in their organizational settings while simultaneously learning from the intervention and assessing it concurrently (Sein et al., 2011, p. 37-38). ADR is seen as applicable when the establishment of an" in-

depth understanding of the artifact–context relationship is needed to develop a socio-technical design agenda for a specific class of problems." (Sein et al., 2011, p. 52-53)

ADR combines ideas from two known research methods: action research (AR) and design research (DR) (Tiainen et al., 2015, p. 19). In AR, the researcher aims to solve practical issues or improvement demands by intervening (changing practice) in an organization. The intervention is done closely with the organization. Results from AR benefit both the organization in the form of problem-solving and the scientific community in the form of new practical knowledge on the subject matter. (Tiainen et al., 2015, p. 2) In DR, the researcher strives to create an IT artifact to solve a problem within the organization (Tiainen et al., 2015, p. 3). However, in DR, as opposed to AR, the researcher is solely responsible for the IT artifact, and organizational context and collaboration do not play a significant role in creating the artifact (Sein et al., 2011, p. 38).

ADR was proposed as a new research method by Sein et al. to address the need for a research method that would recognize more profoundly the organizational context and its effects on the IT artifact. In ADR, the research focuses on creating a suitable IT artifact for the organization as opposed to AR, in which the focus is on making changes to the organization and its activities. (Sein et al., 2011, p. 38-40) However, the precise definition of an IT artifact is still a matter of dispute due to inconsistencies in the term's usage. (Alter, 2015, p. 48-50; Sein et al., 2011, p. 38)

Sein et al. (2011, p. 38-39) view an IT artifact as an ensemble where the organizational domain is structurally engraved into the artifact during development and later usage. Consequently, in ADR, the IT artifact is viewed as an ensemble emerging from the inter-section of development and intent of the researcher, contextual factors, refinement, and usage, as well as its influences on the IT artifact.

The ADR method deals with two challenges: (1) addressing a context-specified problem by intervening and evaluating; and (2) creating and evaluating an "IT artifact that

addresses the class of problems typified by the encountered situation." To fill the requirements of both challenges, the method focuses on building, intervention, and evaluation of the created artifact. The created artifact "reflects on theory, the intent of the researchers, the influence of users and ongoing use in context." (Sein et al., 2011, p. 40)

As the nature of the artifact is an ensemble, it must be stated that ADR deals with the following critical issues:

- Evaluation and building of the artifact are done in cycles and not in sequences as in DR.
- Evaluation of the artifacts should occur naturally and whenever possible, as controlled assessment is challenging to design and conduct.
- "Innovation must be defined for the class of systems typified by the ensemble artifact." (Sein et al., 2011, p. 43-44)

The ADR methods stages and principles addressing the aforementioned issues are seen in figure 6 (adapted from Sein et al., 2011, p. 41).
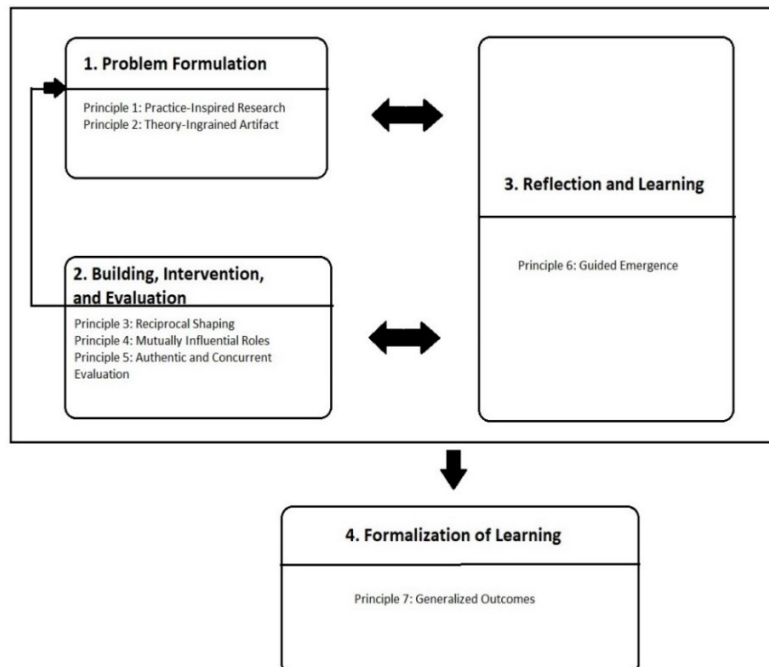
Figure 6. The ADR Method

4.2.1. Problem formulation

The ADR method is triggered by a practical problem or the researcher's initiative. The preliminary empirical investigation aims to "identify and conceptualize a research opportunity based on existing theories and technologies." In addition, the scope, roles in the research, practitioner's participation in the problem solving, and initial research questions are formed in stage 1. (Sein et al., 2011, p. 40)

Two critical elements are identified in the first stage of ADR: ensuring long-term commitment from the participating organization and defining the problem as a class of problems. The anchoring principles in stage 1:

- Principle 1: Practice-Inspired Research.
- Principle 2: Theory-Ingrained Artifact.

(Sein et al., 2011, p. 40)

Principle 1 views practical problems as "knowledge creation opportunities" at the organizational domain and technology intersection. Action design researcher is expected to generate knowledge that applies to a class of problems exemplified by the organization's problem. Therefore, the ADR team is not expected to solve the practical problem in the organization rigorously but to merely "intervene within the organizational context of the problem." (Sein et al., 2011, p. 40)

Principle 2 requires that the artifact is based on theory. The initially designed artifact should be found on the generalized theory where the researcher is inscribing theoretical elements into it. Afterward, the artifact is subjected to "cycles of intervention, evaluation and reshaping" in the organizational context. (Sein et al., 2011, p. 40-41)

4.2.2. Building, intervention, and evaluation

Stage 2 is based on the framed problem and theoretical premises from stage 1, which works as a groundwork for the initial design of the IT artifact. Subsequent development phases take place in stage 2. (Sein et al., 2011, p. 41)

Artifact building, organizational intervention, and evaluation (BIE) are interwovenly carried out as an iterative process in the organizational context. Constant assessment of the problem and the artifact and the articulation of design principles occur in the BIE stage. The result of the BIE stage is the realized artifact. The BIE stage also dictates the whereabouts of the innovation: innovation from the design of the artifact or the organizational intervention. (Sein et al., 2011, p. 41-42)

Stage 2 has two endpoints for the research design continuum: IT-dominant BIE and organization-dominant BIE. At one end, the IT-dominant BIE focuses on creating an innovative technological design. The more mature version of the artifact (beta version) is introduced in the organizational setting. Subsequently, a new BIE cycle is started, or the researcher exits the project. (Sein et al., 2011, p. 42)

On the other end is the organization-dominant BIE most convenient for generating design knowledge primarily from the intervention. In the organization-dominant BIE, the ADR team challenges existing ideas and assumptions about the artifact's usage in the context. In this BIE, the artifact is introduced in the organizational setting in an earlier phase (alpha version). (Sein et al., 2011, p. 43)

Stage 2 has three principles with emphasis on the inseparability of the influencing domains to the artifact:
- Principle 3: Reciprocal Shaping.
- Principle 4: Mutually Influential Roles.
- Principle 5: Authentic and Concurrent Evaluation.

(Sein et al., 2011, p. 43)

Principle 3 underlines the strong inseparable influences from the IT artifact and the organizational context. Recursive cycles make it possible to gain an increased understanding of the organizational context and, therefore, adjust its interpretation and change the chosen design constructs if needed. (Sein et al., 2011, p. 43)

Principle 4 focuses on mutual learning. While researchers offer theoretical knowledge, the practitioners offer insights into organizational presumptions and policies. Contributions from different participants might compete with or complement one another. Individual participants could have multiple roles. However, clarity of assignment responsibilities is worth pursuing the sake of the research experience. (Sein et al., 2011, p. 43)

Principle 5 accentuates that evaluation is an ongoing and interwoven part of the research process and not a subsequent stage. Evaluation cycles for beta and alpha versions differ from one another. "Evaluation cycles for the alpha version are formative, contributing to the refinement of the artifact and surfacing anticipated and unanticipated consequences." As opposed to the evaluation for the more mature beta version assesses value and utility outcomes. Authentic evaluation is seen as more prolific than the hard to engineer controlled evaluation. (Sein et al., 2011, p. 44)

The main differences between IT-dominant BIE and organization dominant BIE are showcased in figure 7 (Adapted from Sein et al., 2011, p. 42-43).
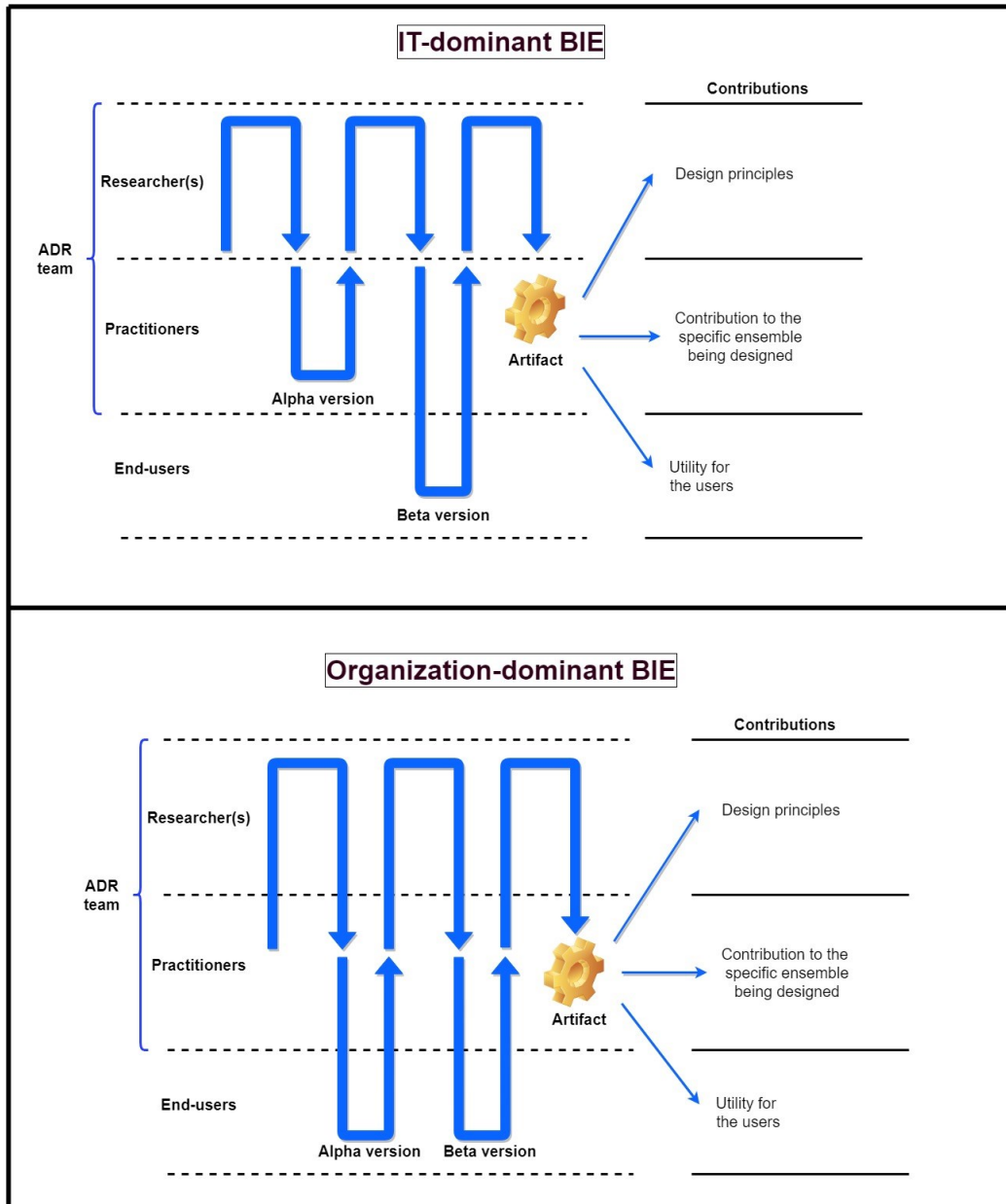
Figure 7. IT-dominant BIE vs. Organization-dominant BIE

4.2.3. Reflection and learning

In stage 3, the research moves from solution building to a specific instance into an extensive class of problems. Stage 3 is a "continuous stage and parallels the first two stages as depicted" in figure 6. (Sein et al., 2011, p. 44)

Stage 3 reflects the research process and sees it as more than simple problem-solving. To ensure that knowledge is genuinely identified, conscious reflection on the theories, problem framing, and the emerging ensemble is required. Based on early evaluation results, adjustments might be necessary to understand the artifact research process better. (Sein et al., 2011, p. 44)

The only principle in stage 3 is guided emergence (Principle 6). Contradicting terms emphasize the dynamic reflection of the ensemble artifact on the initial design, current development in the organizational context, and simultaneous outcomes from authentic evaluation. Rising sensitive signals implying the need for trivial and substantial refinements are expected to be dealt with in a project. (Sein et al., 2011, p. 44)

4.2.4. Formalization of learning

According to Sein et al. (2011, p. 44), the fourth stage is focused on formalizing the learning. "Learning from problem-specific solutions should be transformed into solution concepts for a class of field problems." The outcomes can be turned into design principles refining theories that influenced the initial design. Stage 4 draws from the generalized outcomes principle (Principle 7). Due to the situated nature of ADR, generalization is seen as challenging to achieve. The ensemble artifact "represents a *solution* that addresses a problem. " In ADR, the transition "from specific-and-unique to generic-and-abstract" is critical. A three-level conceptual move to address this is suggested:

1. Problem instance generalization
2. Solution instance generalization
3. Derivation of design principles

Principle 7 deals with casting known factors as instances of their classes. The problem, the solution, and design principles (knowledge capturing) are cast into their respective classes. Finally, through derivation, it is possible to connect generalized outcomes, the

design principles "to a class of solutions and a class of problems." (Sein et al., 2011, p. 45)

Figure 8 (based on figure 1 by Sein et al., 2011. p. 41) is an adapted and modified figure that presents the ADR as a concept, including tasks related to each phase.
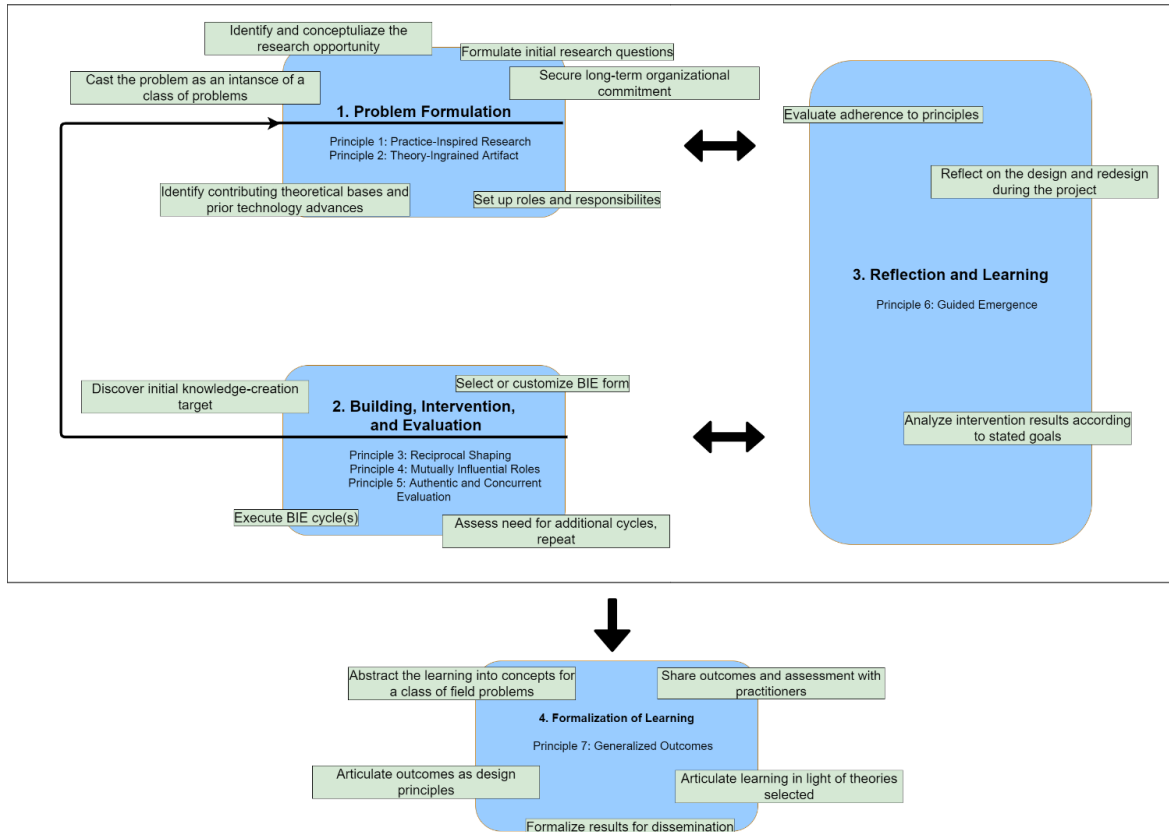
Figure 8. ADR Method including associated tasks

4.3. Data collected for the study

The study data collected consisted of private limited liability companies in Finland. Of those, a specific form of a customer group was chosen as the focus of the artifact: LLCs with estimated tax years. In this study, financial data/information was considered from the years 2017 to 2019 of the companies conducting solely business activities under the act on business income only. The companies with personal and agricultural income

sources were excluded. Data for tax years 2017, 2018, and 2019 consisted of information from 94 889 companies in the original dataset and 203 617 companies in the corrected dataset. Data gathered from the tax system included specific information about the companies per tax year, as presented in table 4.

Table 4. Dataset description

1.	The main line of business as a 2-digit code expressing to which specific business sector the company belongs
2.	Starting date in taxation
3.	Starting date in the trade register
4.	Closing date in taxation *
5.	Closing date in trade register *
6.	Reason for closing the trade register *
7.	Home municipality
8.	Tax year (2017, 2018, or 2019) **
9.	Net sales per year (€)
10.	The total taxable business income per year (€)
11.	Purchases, variation in stocks and inventory per year (€)
12.	Total tax-deductible business costs per year (€)
13.	Assets total per year (€)
14.	Liabilities total per year (€)
15.	Taxation estimated at some point (yes/no)
16.	Taxation still estimated (yes/no)

* = information expressed if available

** = tax year is defined by the closing date of the accounting period, i.e., the accounting period ends 31.1.2018, the tax year is 2018

Standard Industrial Classification TOL 2008

The main line of business is based on Standard Industrial Classification TOL 2008, which is formed from five hierarchical levels. This thesis utilizes only the first two levels. (Statistics Finland, 2021)

Closing a limited liability company

A limited liability company can be closed (dissolved) by going into liquidation by decision of the General Meeting, through a merger or demerger, bankruptcy, deregistration, or liquidation by order of the authority. (Finnish Patent and Registration Office, 2014)

5 Developing a knowledge creation artifact

This chapter presents how ADR was utilized to form the IT artifact(s), how the creation proceeded, how the IT artifact performed, what was learned, and what kind of design principles came out of the process. This project aimed to create a suitable IT artifact for the case organization, the Finnish Tax Administration, according to the ADR research methods definition presented in chapter 4.

5.1. Problem formulation

Building an IT artifact to address the identification problem on the claim for adjustment procedure rose from the researcher's interest in machine learning and its newest advances in taxation and fields similar to it. Such studies include but are not limited to Chen et al. (2011) with an automatic detection model for tax reports, Xiangyu et al. (2018) with a NN focusing on tax evasion, and Pérez López et al. (2019) with a NN identifying tax fraudulent persons. Additionally, Zhang et al. (2020) developed a novel approach, with NN identifying tax evasion from social media posts. The studies indicate that the technology has been increasingly applied in practice with encouraging outcomes. (Sein et al., 2011, p. 40)

As the researcher was an employee processing the claims in the LLC's claims of adjustment procedure, he had a piece of firsthand knowledge of the problem in its context. The binary nature of the problem (has been estimated or not) and the actual need to decrease the amount of processing time in the claim for adjustment procedure resulted in the researcher presenting an idea to create an IT artifact based on AI algorithms to tackle this issue.

The presented idea of an IT artifact was met with interest in the case organization. The creation of such an artifact was seen as beneficial in multiple ways. It could potentially

increase knowledge of the problem itself, offer solutions to the problem, and increase understanding of the applicability of the said technology. Long-term commitment from both the researcher and the organization was secured with a contract.

The process of carrying out taxation is expected to be performed according to the law. Since taxation is mainly automated, it has raised questions about what can be and should carry out automatically. The deputy ombudsman pointed out in 2019 that automation in taxation is not based on precise legislation, and it should only be applied to tax assessment but not to decision-making. This current topic and policy set limitations for the artifact and its functionality.

The ethical principles for AI usage were followed in creating the artifact as well as possible. The data used was monitored and analyzed, and corrections were made when necessary. The artifact was not an independent decision-maker as a human monitored its performance and did not function as a decision-maker. Although legal and ethical perspectives were not the focus of this research, they provided an essential and holistic starting point for creating the artifact.

The problem is presented as a binary classification problem representing a class of tax administration problems. The class represents problems that require identification. By recognizing the instances from one another and identifying underlying issues, the organization could pre-emptively decrease the occurrence of these problems.

Roles and responsibilities in the ADR team were set at the beginning of the project. The researcher was responsible for the research and creating the IT artifact. An analytics expert and claims of adjustment procedure representative offered valuable guidance and comments on the problem and the data for the project.

To create such an artifact, the researcher was handed intuitively chosen and anonymous data on LLCs from tax years 2017 to 2019. Intuitive data selection was performed by

professionals working in the problem area who have developed a thorough understanding of the problem in its context. Participating in data selection were the researcher, analytics expert, and claims of adjustment procedure representative, who also formed the ADR team. Figure 9 depicts the different viewpoints of the team members in the project.

A: Researcher

-Research
-Artifact building
-Firsthand knowledge of the problem domain

B: Process representative

-Expert knowledge of the problems in the process
-Artifact evaluation

C: Analytics expert

-Data retrieval process
-Expert knowledge in analytics
-Artifact evaluation
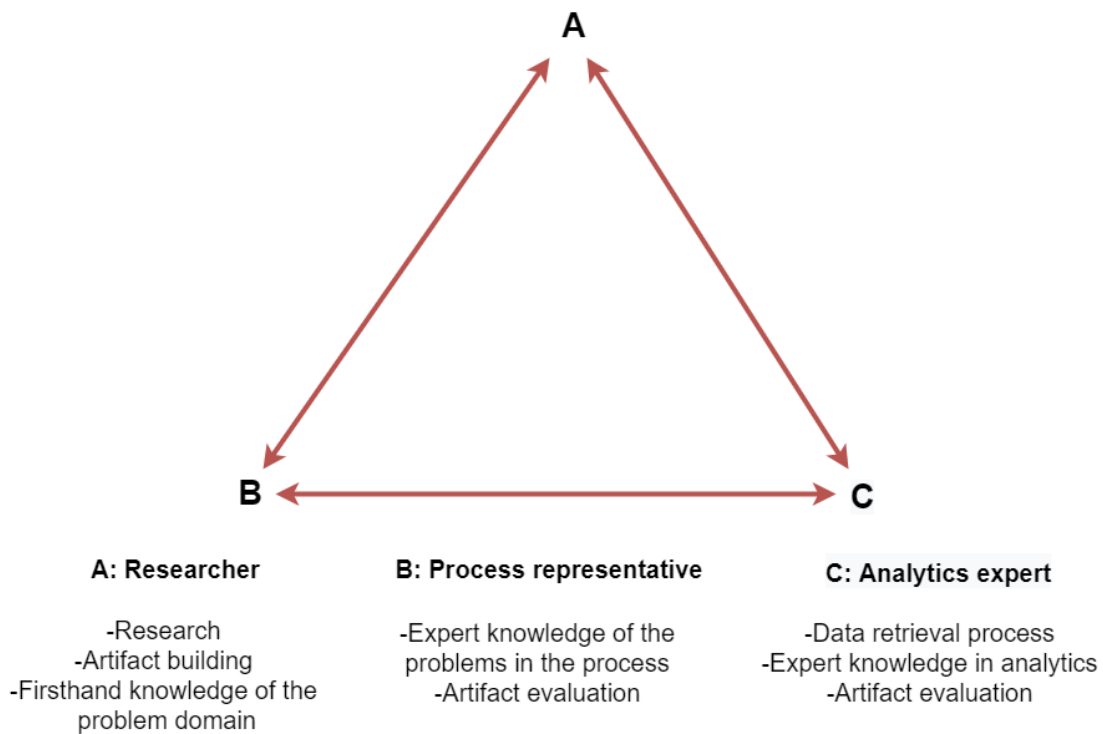
Figure 9. BIE viewpoints in the project.

The IT artifact was not intended as a decision-making instrument but as a knowledge creation instrument due to legal and ethical limitations to what an automated solution is allowed to do and what is expected from an AI solution within the problem context. The information provided by the artifact would only be used proactively. Within this problem context, it could remind potential customers who are prone to not returning their tax returns according to the information provided by the artifact. Reminding could occur, for example, as a guidance text message or as an instruction letter.

IT-dominant BIE-cycle was chosen as the focus of this ADR project. The researcher and the organization were interested in how well the artifact could perform as a knowledge-

creating pre-emptive instrument. Restrictions stemming from law and ethics were considered in data retrieval and potential artifact usage contexts. The artifacts beta version was not introduced to the end-users as it was not possible within the timeframe.

The IT artifact manufactured was not meant to be handed over to the organization. The artifact was tested on the data received from the organization, and its performance was analyzed with the ADR team. Figure 10 presents the blueprint for the BIE plan in the project.
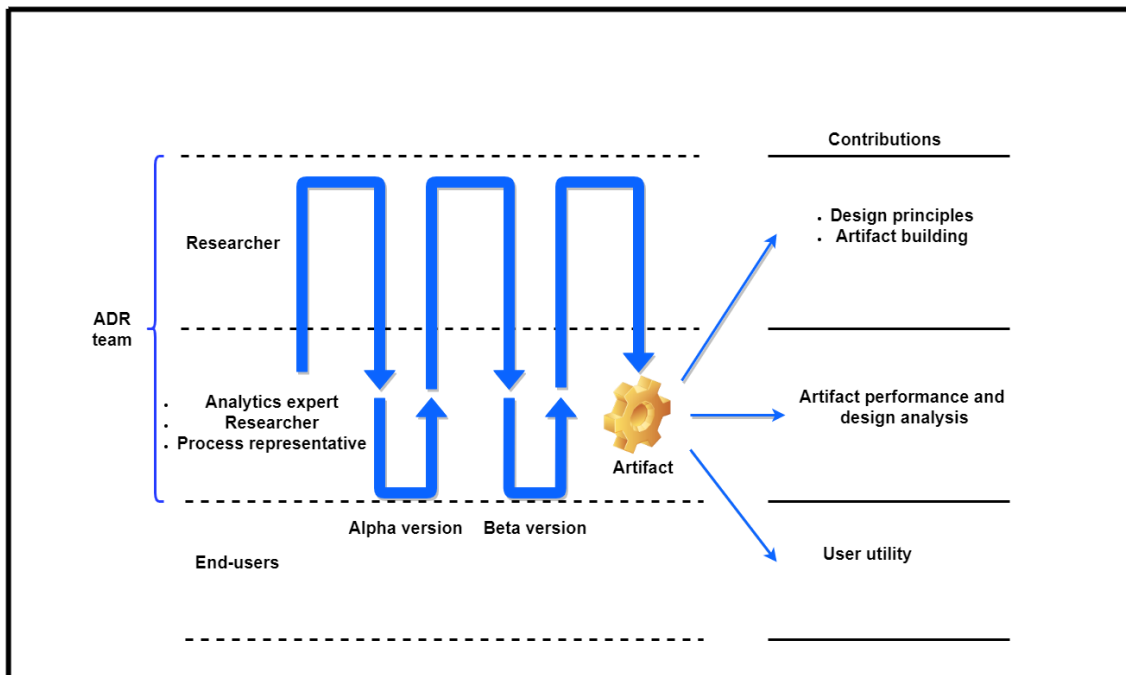
Figure 10. Original IT-dominant BIE plan in the tax administration ADR project

5.2. Building, intervention, evaluation

The ADR team analyzed the problem and decided on which data to retrieve for the project. Data consisted of information presented in tax returns and basic information regarding the companies. Information such as names, owners, and company identification numbers was excluded. After being handed the data, the researcher began creating the artifact. The artifact aimed to recognize the estimated companies from non-estimated

ones as efficiently as possible. The initial knowledge creation target was to determine if such technologies could potentially be used for taxation benefits. The performance was monitored with precision, recall, f1-score, and confusion matrix.

IT artifact in this project was a piece of code analyzing data with DL libraries to deduce a correct outcome (supervised learning). The creation of the IT artifact started from scratch and was mainly based on trial and error with Keras. The characteristics of the NN were modified and tested in cycles. The data available was trimmed down to find out the variables that had the most substantial impact on the result.

Keras (2021a) is a DL application programming interface written in Python on the ML platform Tensorflow. The code was written with Spyder IDE (Spyder, 2021), a scientific Python development environment. Strong theoretical background in machine learning and neural networks' ability to enhance processes worked as the starting point for this project. The first functional version of the IT artifact, the alpha version, and its results was analyzed with the project team. The results are presented in tables 5 to 7 and figures 11 to 13.

5.2.1. Alpha

Table 5 presents the results for the alpha performance with tax years 2017 test set. Precision, recall, and f1-score are calculated for the non-estimated companies (0) and estimated companies (1). The number of companies in each group is presented under "Number of companies." The macro average refers to the average of the NN's performance in both classes. Tables 6 and 7 are structured the same way as table 5 way. They present the performance results attained with the artifact created with 2017 test data.

Table 5. Alpha, performance, 2017 test set

Estimation sta-tus	Precision	Recall	F1-score	Number of compa-nies
0	0.97	0.98	0.98	26822
1	0.61	0.55	0.58	1645
macro average	0.79	0.76	0.78	28467

Figure 11 presents the CM for alphas performance on 2017 test set. As an example, precision (TP / TP + FP) for class (estimation status) 1 is calculated: 905 / (905 + 569) = 0.61. Figures 12 and 13 presents how the artifact that was create with 2017 test data performed with 2018 and 2019 data.

Figure 11. Alpha, CM, 2017 test set

Table 6. Alpha, performance, 2018

Estimation sta-tus	Precision	Recall	F1-score	Number of compa-nies
0	0.98	0.98	0.98	93062
1	0.58	0.57	0.58	4920
macro average	0.78	0.77	0.78	97982

Figure 12. Alpha, CM, 2018

Table 7. Alpha, performance, 2019

Estimation sta-tus	Precision	Recall	F1-score	Number of compa-nies
0	0.98	0.98	0.98	96388
1	0.48	0.56	0.52	3677
macro average	0.73	0.77	0.75	100065

Figure 13. Alpha, CM, 2019

As the data was highly imbalanced, it proved to be an arduous task to get the neural network to recognize the estimated companies' class (1) as well as possible. The precision, recall, and f1-score for the 0-class were excellent in the alpha-version. However, the estimated class (1) results were not on an acceptable level. The neural network's f1-score for the estimated class achieved 0.52-0.58, which was not on par with the f1-score of the 0-class's 0.98. Reaching at least a moderate f1-score was set as a target for the beta version.

The project team analyzed the performance of the alpha version and concluded that further refinements were required, and new approaches were suggested as well as different testing scenarios. These worked as the starting point for the development of beta-version. The researcher continued the development according to the original BIE cycle.

5.2.2. Beta

The researcher selected three different approaches to be tested in creating the beta version. Alpha-version worked as a starting point for the beta. The test plan for the beta version can be seen in table 8:

Table 8. Project team's test plan for the beta

Test approach	Results
Creation of new artificial variables from data	<u>Variable 1</u> Is the company passive and empty (no assets, liabilities, sales, purchases, or other taxable activity). Yes (1) or no (0). Using only variable 1, the neural network achieved similar results as the alpha version. <u>Variable 2</u> Is the company passive but has 10 000 or more in assets? Yes (1) or no (0). This variable did not affect the performance, and acceptable results were not achieved using only this variable or having it as an additional variable.
Creating the artifact from the tax year 2018's data and tested on 2019 (leaving 2017 out altogether)	By leaving out data from the tax year 2017, no improvements were achieved nor significant drops in performance.
Eliminating unnecessary variables from the data	By eliminating variables, a slightly better performance was reached. The remaining variables: • total taxable business income, • total tax-deductible business costs, and • net worth Results are shown in tables 9 to 11.

The results of the beta version were analyzed with the ADR team. Minor improvements to the model's performance were achieved compared to alpha. Tables 9, 10, and 11 present the result the same way as the results for the alpha version were previously presented.

Table 9. Beta, performance, 2017 test set

Estimation status	Precision	Recall	F1-score	Number of companies
0	0.98	0.98	0.98	17909
1	0.62	0.59	0.60	1069
macro average	0.80	0.78	0.79	18978

Table 10. Beta, performance, 2018

Estimation sta-tus	Precision	Recall	F1-score	Number of compa-nies
0	0.98	0.98	0.98	93062
1	0.58	0.60	0.59	4920
macro average	0.78	0.79	0.78	97982

Table 11. Beta, performance, 2019

Estimation sta-tus	Precision	Recall	F1-score	Number of compa-nies
0	0.99	0.97	0.98	96388
1	0.49	0.64	0.55	3677
macro average	0.74	0.81	0.77	100065

The increased performance in recognizing the estimated companies (class 1) can be seen in the increased f1-score from alpha to beta in all the tax years from 2017 to 2019. Table 12 presents how f1-score increased during the artifact development from alpha to beta. F1-score %-increase refers to the relative increase in performance from alpha to beta.

Table 12. Artifact's performance increased from alpha to beta

Version	Tax year	Precision	Recall	f1-score	f1-score %-increase
Alpha	2017	0.61	0.55	0.58	-
Beta	2017	0.62	0.59	0.60	3,5%
Alpha	2018	0.58	0.57	0.58	-
Beta	2018	0.58	0.60	0.59	1,7%
Alpha	2019	0.48	0.56	0.52	-
Beta	2019	0.49	0.64	0.55	5,77%

The most encouraging finding was that the model's performance was maintained and slightly increased by eliminating variables. Throughout the testing phases in alpha and beta versions, the NN recognized the companies that were not estimated with an f1-score of 98%. It was expected because the data was strongly imbalanced towards the non-estimated class. Moreover, as pointed out by the claim of adjustment procedure representative, it has always been a challenge to recognize the estimated companies from others in practice.

During the testing phase of the beta version, it was discovered that a problem had occurred in the data retrieval process. As a result, fifty percent of the data had been missing. It was decided that no additional development cycles would take place, but instead, the finalized artifact would be tested with the corrected data that now included double the amount of data. However, the relative size of the estimated class did not change, and the dataset was still largely imbalanced between the classes.

Tables 13 to 15 and figures 14 to 16 present the results with the beta version of the artifact (no parameters changed) but utilizing the corrected dataset. Results are presented in the same way as in alpha and beta.

Table 13. Beta with additional data, performance, 2017 test set

Estimation status	Precision	Recall	F1-score	Number of companies
0	0.97	0.98	0.98	37320
1	0.64	0.57	0.60	2208
macro average	0.81	0.77	0.79	39528

Figure 14. Beta with additional data, CM, 2017 test set

Table 14. Beta with additional data, performance, 2019

Estimation status	Precision	Recall	F1-score	Number of companies
0	0.98	0.98	0.98	193477
1	0.59	0.59	0.59	10140
macro average	0.78	0.79	0.79	203617

Figure 15. Beta with additional data, CM, 2018

Table 15. Beta with additional data, performance, 2019

Estimation status	Precision	Recall	F1-score	Number of companies
0	0.99	0.98	0.98	200284
1	0.51	0.63	0.56	7580
macro average	0.75	0.80	0.77	207864

Figure 16. Beta with additional data, CM, 2019

Table 16. Artifact performance comparison

Version	Tax year	Precision	Recall	f1-score	f1-score %-increase
Beta	2017	0.62	0.59	0.60	-
Beta (corrected data)	2017	0.64	0.57	0.60	0%
Beta	2018	0.58	0.60	0.59	-
Beta (corrected data)	2018	0.59	0.59	0.59	0%
Beta	2019	0.49	0.64	0.55	-
Beta (corrected data)	2019	0.51	0.63	0.56	1,8%

No significant improvements were achieved even though the amount of data was doubled, as seen in table 16. As an additional point of view for analyzing the problem, the analytics expert suggested that analyzing the data with a self-organized maps algorithm could provide valuable insight into the problem domain from UL's perspective. It was decided that another development cycle (gamma) would be conducted by testing the performance of a SOM algorithm on the data. Therefore, changes to the original BIE plan were made. Changes and actualized BIE cycles are presented in figure 20.

5.2.3. Gamma

Based on the meeting regarding the beta version, an additional cycle was conducted that would utilize a UL method to tackle the issue of recognizing estimated companies. SOM was chosen as the UL method. The artifact was built with a SOM library (Minisom, 2022) created by Vettigli (2018).

In figures 17 to 19, five points representing the green and red squares and the white/green points were selected. Green squares represent points where there are no estimated tax years. A white point indicates that the data differentiates from the rest. Red circles represent companies whose tax year was estimated. A point including red and green indicates that a clear distinction between the two was not attained. The income, expenses, and net worth values represent the median of the values in the specific points selected. Numbers in the X and Y axes represent the coordinates of each point. For example, in figure 17, the white point with a green circle has coordinates x=11 and y=3. The mentioned points specific company-related information in median values is presented in table 17.
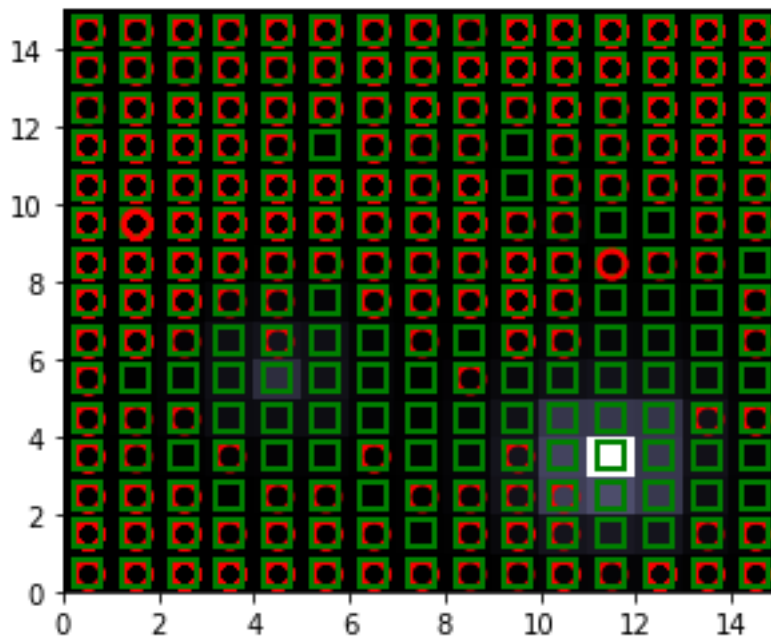
Figure 17. SOM, 2017

Table 17. Datapoints, 2017

Point	Coord.	Income	Expenses	Net worth	Firms
G	x=2, y=5	75 379 €	64 361 €	20 005 409 €	77
G	x=5, y=5	641 946 €	2 117 371 €	142 718 888 €	19
G	x=12, y=6	12 855 435 €	11 769 776 €	9 159 407 €	86
G	x=5, y=11	2 205 864 €	1 563 443 €	1 661 337 €	178
G	x=9, y=11	1 751 034 €	1 571 442 €	674 004 €	505
R	x=1, y=9	171,00 €	974,00 €	24 944,00 €	5072
R	x=11, y=8	3 137 743,00 €	2 800 569,00 €	1 534 862,00 €	169
R	x=14, y=14	- €	- €	116,00 €	30737
R	x=13, y=12	52 574,00 €	51 841,00 €	4 936,00 €	6915
R	x=4, y=9	223 652,00 €	212 214,00 €	29 866,00 €	5123
W	x=11, y=3	825 686 950,00 €	748 791 489,00 €	163 640 054,00 €	76

Based on the SOM map of the tax year 2017 (figure 17; table 17) and the points selected, it is evident that the SOM algorithm does perform well at recognizing the non-estimated class. Companies with high income and expenses and significant net worth fall in the green area. The area containing white represents part of the data that differs from the rest. In this case, the 76 firms in the area are most likely the largest and most active

companies in Finland regarding income, expenses, and net worth. The estimated class is not distinguishable. However, the areas that include red are lesser in income and expenses and more minor in net worth. The largest number of firms fall in point (14, 14) where the median income, expenses, and net worth are low.
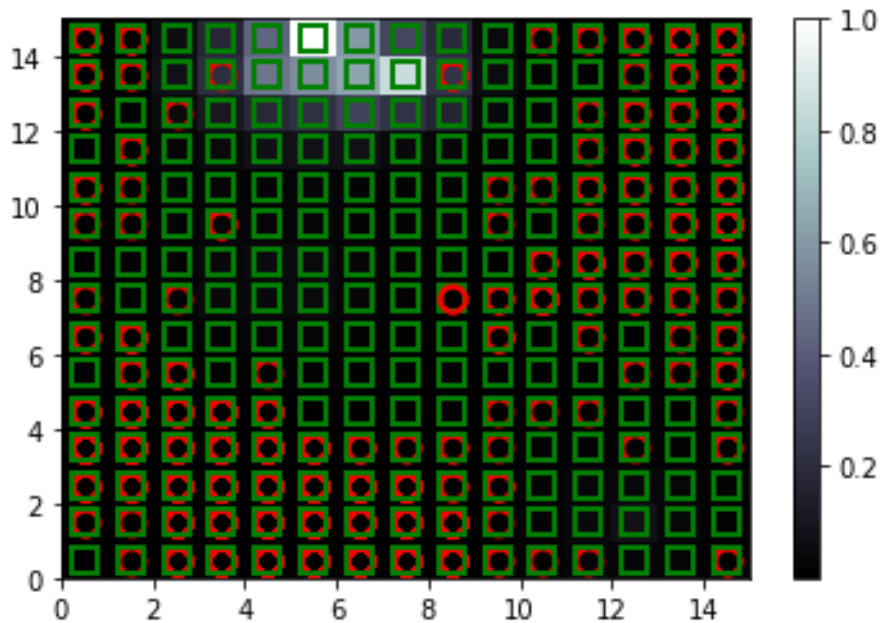

Figure 18. SOM, 2018

Table 18. Datapoints, 2018

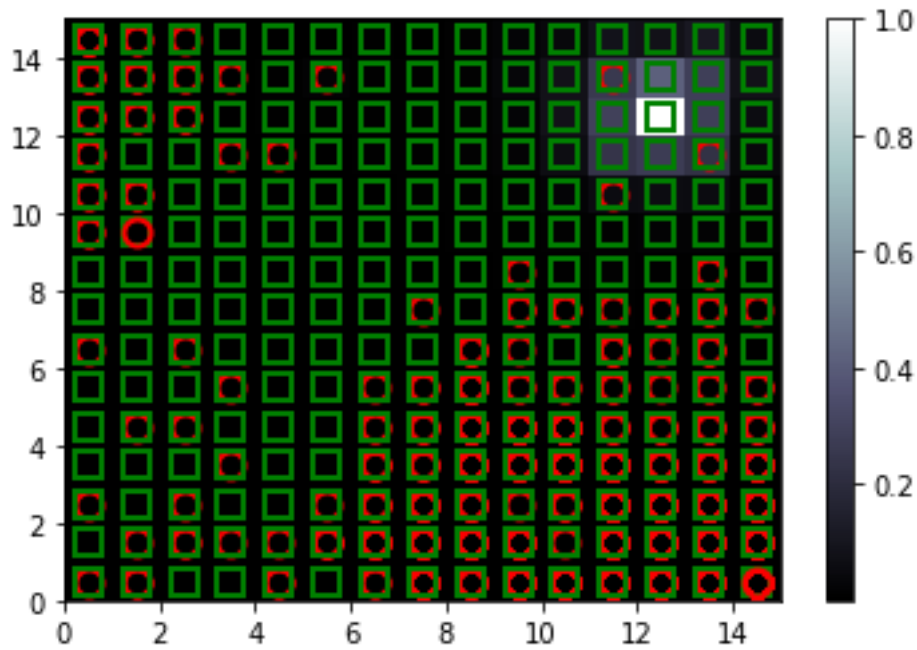| Point	Coord.	Income	Expenses	Net worth	Firms
G	x=5, y=11	49 867 261,00 €	47 189 423,00 €	12 304 960,00 €	24
G	x=9, y=12	12 249 065,00 €	11 009 212,00 €	11 809 686,00 €	28
G	x=12, y=2	15 426 474,00 €	10 993 777,00 €	53 671 393,00 €	34
G	x=4, y=8	65 621 127,00 €	64 488 262,00 €	8 151 510,00 €	161
G	x=10, y=6	1 701 876,00 €	1 520 295,00 €	613 394,00 €	468
R	x=6, y=1	36 290,00 €	35 013,00 €	6 693,00 €	12708
R	x=3, y=0	1 639,00 €	3 061,00 €	167 178,00 €	3356
R	x=14, y=14	296 326,00 €	353 774,00 € -	164 577,00 €	1027
R	x=8, y=7	5 060 955,00 €	4 592 170,00 €	2 161 655,00 €	120
R	x=5, y=0	- €	514,00 €	900,00 €	47328
W	x=5, y=14	1 453 362 342,00 €	1 442 106 548,00 €	246 206 271,00 €	42
W	x=7, y=13	20 411 718,00 €	5 592 390,00 €	933 404 590,00 €	36

Figure 19. SOM, 2019

Table 19. Datapoints, 2019

Point	Coord.	Income	Expenses	Net worth	Firms
G	x=12, y=9	20 802 743,00 €	20 321 578,00 €	2 493 762,00 €	176
G	x=11, y=8	13 804 352,00 €	13 442 689,00 €	1 511 917,00 €	236
G	x=14, y=11	59 812 933,00 €	59 114 367,00 €	8 017 102,00 €	154
G	x=14, y=8	6 515 351,00 €	6 395 230,00 €	1 859 024,00 €	38
G	x=5, y=12	4 409 458,00 €	3 795 789,00 €	43 377 505,00 €	34
R	x=1, y=9	1 527 105,00 €	1 314 091,00 €	2 234 482,00 €	191
R	x=8, y=4	1 060 912,00 €	1 047 197,00 €	68 098,00 €	1391
R	x=13, y=1	44 207,00 €	41 660,00 €	9 563,00 €	12120
R	x=14, y=1	16 894,00 €	16 563,00 €	9 210,00 €	15569
R	x=14, y=0	- €	264,00 €	1 276,00 €	48569
W	x=12, y=12	1 465 951 470,00 €	1 435 013 955,00 €	309 162 510,00 €	41

The SOM map performs similarly with tax years 2018 and 2019 data (figures 18 and 19). More active and wealthy companies fall into the non-estimated class, and smaller and non-active tend to fall into the red/green areas (tables 18 and 19). The white areas are prone to represent the largest and most active companies. Areas including red are lesser in income and expenses and more minor in net worth. The largest number of firms for

the tax year 2018 fall in point (5, 0), where the median values of income, expenses, and net worth is low. For the tax year 2019, this is point (14, 0). A clear distinction between whether a company with lesser income, expenses net worth type falls into an estimated class is not attained.

SOM algorithm performed similarly to the beta version. The non-estimated class was detected more efficiently than the estimated class. SOM algorithm provided information that narrows down where the estimated companies are more likely to occur.
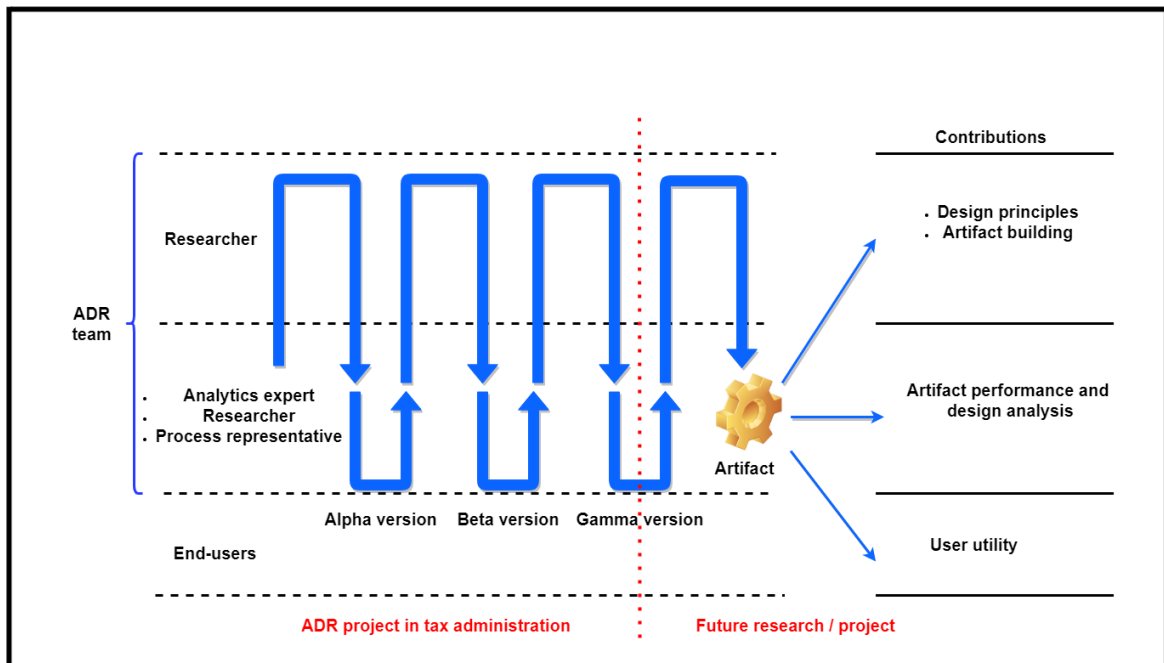
Figure 20. Actualized IT-dominant BIE cycles in the project

Due to time-related restraints, the BIE cycle was concluded with the gamma version (figure 20). Multiple additional cycles could include either continuing the development of beta and gamma or creating an entirely new artifact. It is beneficial to develop and test more than one potentially suitable artifact. Results from beta and gamma provide information that an AI artifact has potential as a pre-emptive, knowledge-creating analysis instrument. Moreover, information created by two different algorithms strengthens and ratifies one another. End-user participation and testing in practice were absent from the

project as they were not possible to conduct. In addition, they are left for future research and projects that would advance artifacts of which development began in this project. This project's AI-empowered analysis algorithms, such as the beta and gamma versions, are potential tools for decreasing irregular taxation-related behavior. (Sein et al., 2011, p. 44)

5.3. Reflection and learning

The artifact creation followed the principles laid out by ADR, considering project-related restrictions in time and scope. The researcher's dual roles (the research itself and the artifact's creation) slowed down the process. Having another person responsible for the actual artifact creation would have benefited the project. The project would have required more focus on setting up roles and responsibilities. However, the communication within the project team functioned well. The artifact's creation followed the IT-dominant BIE, and the whole project was continuously evaluated (stages 1-3). As a result of the reassessment conducted after the beta versions evaluation, a modification to the original BIE plan was added during the project.

The concept for the artifact stemmed from theory (theory-based artifact), and its creation was motivated by practice, thus creating a link between the two. Concurrent and authentic analysis of the artifact's performance was vital for the process. Everybody involved in the project could suggest and affect the direction of the artifact development.

The object of the project was to create an IT artifact based on AI capable of identifying companies not returning their tax returns. After being handed the data, the researcher started creating the code in Python, following Keras guidelines. It was decided to analyze the model's performance with a confusion matrix, precision, recall, and F1-score. The initial artifacts creation ended with the beta version.

The development process from alpha to beta and gamma benefited from interventions as insightful ideas to test out helped shape the form of the artifact. As a result, from an intervention, the idea of creating an entirely different artifact for the same problem was initialized, resulting in the artifact's gamma version. The change emerged from the research methods guided emergence principle, which expects the project team to be sensitive toward expected and unexpected consequences and act on refinement needs even though substantial changes would occur.

The Gamma version approached the problem by analyzing the data with a SOM algorithm instead of the original vision. Gamma achieved similar results to beta by recognizing non-estimated companies and having issues drawing a distinct line between estimated and some non-estimated companies. The creation of gamma pinpoints the need to approach problems with more than one AI algorithm allowing users to attain knowledge from several different algorithms that build on one another.

The project followed ADR principles according to the IT dominant BIE with certain limitations. End-user testing was left out since the project was concluded with a gamma version. More development would be required to reach a finalized artifact that could be deployed to end-users.

5.4. Formalization of learning

The fourth stage in the ADR project required a change from specifics to generalization, divided into three levels according to the research method. The problem that ignited the project is that of a default recognition utilizing historical data (a generalization of the problem instance).

The presented solution to the generalized problem is an AI-empowered instrument that is fed historical data and, based on that, attempts to classify the customers into defaulting and non-defaulting ones. This insight provided by the artifact is used in decision-

making. The knowledge-creating instrument offers a basis for pre-emptive actions as suggested in this project (a generalization of the solution instance). (Sein et al., 2011, p. 45)

The design principles were formed based on the answers to the research questions and how the artifact performed in practice. The research questions answered:

1. How can AI be deployed to the case organization to create value in its current taxation system?

Value is created in the form of time, speed, and liberated resources. AI-empowered instruments can analyze and form inferences significantly faster than a human could. Therefore such a solution would free time for human(s) to concentrate on more urgent matters. It is left for human(s) to analyze the results provided by AI to see if it is applicable. The ADR project in this thesis proved that a NN and a SOM algorithm could detect estimated (defaulting) companies even though room for improvements was left.

As the SOM analysis proved, AI is also better at detecting patterns and segmenting customers into corresponding sectors. Finland's most valuable and active companies are likely not among the SOM map's estimated (defaulting) companies area. Added value can be created by utilizing AI to help decision-making if boundaries on where and how to use the information are well examined. Moreover, it is required that AI does not function as a decision-maker, and a human is always behind every tax-related decision.

The information provided by AI in the problem context either provides specific information on companies within a particular sector as in the SOM analysis (descriptive) or information on whether the company will be estimated (imperative). To gain the most out of AI usage, it is suggested to approach a problem with more than one artifact. Doing this makes it possible to attain a more comprehensive view of the problem and its potential solution.

2. How can AI be deployed so that it does not violate rules and regulations?

Challenges in using AI in taxation can be divided into three levels :

 1. Trustworthiness through accuracy

To be utilized as a knowledge creation instrument, adequate performance and accuracy are required and expected. Thorough testing, analysis, and continuous refinement are mandatory. Unless an organization-defined acceptable performance for an AI solution is not achieved, the solution should be discarded, and a new approach should be taken.

2. Legal and ethical restrictions and limitations of use

Legal and ethical perspectives should always be considered when developing an AI solution. According to the Finnish Tax Administration's ethical principles for AI: AI should only use reliable data, follow laws and legislation, and constantly be monitored and managed by a human. Legal and ethical matters need to be addressed and considered when building AI solutions. They set restrictions, expectations, and requirements for an AI solution. A transparent and regulated AI solution is expected.

3. Justification of usage

A preliminary inspection on where to use AI solutions should be conducted to deduce if significant improvements are achievable. Restrictions and limitations of use should always be taken into account. Problems with a lot of data available might make a desirable use case for AI. Organizations should prepare a few different AI approaches to tackle an issue that could potentially be solved entirely or partially with AI. Additionally, as the cost of using AI has decreased, organizations ought to have a low threshold for experimenting with them.

Lastly, seven meetings were held concerning the ADR project. Presentation and evaluation of alpha and beta versions formed the project's core. Outcomes achieved in the

project were shared with the organization, including the gamma version and its findings. Encouraging results pave the way for future projects and the development of similar solutions. Dissemination of results was left out as the ADR process was not finished, and a finalized product was not created.

The preliminary design principles derived from research outcomes achieved until gamma iteration are presented in table 20.

Table 20. A preliminary set of design principles

Design principle	Description
Trustworthiness through accuracy	Trust is achieved only by producing accurate results. Organizations decide which is the acceptable level of accuracy.
Legal and ethical restrictions and limitations of use	Boundaries set by legal and ethical viewpoints should be an integral part of the development from the beginning and continuously evaluated to achieve sustainable and transparent use of AI.
Justification of usage	A preliminary investigation of the problem and possible AI solutions should be undertaken to determine if significant benefits are attainable.

A summary of the ADR process focusing on creating a pre-emptive artifact is shown in table 21. Table 21 is an adapted table based on Sein et al. (2011, p. 51).

Table 21. Summary of the ADR process

Summary of the ADR Process in the pre-emptive artifact for tax project		
Stages and Principles		**Artifact**
Stage 1: Problem Formulation		
Principle 1: Practice-Inspired Research	Practical challenges in the case organization and the willingness to explore novel solutions worked as a launching point.	**Recognition**: Interest in utilizing AI solutions in taxation has grown. A transparent and holistic approach is vital. Simple problems with plenty of data are potential use cases such as recognizing estimated companies.
Principle 2: Theory-Ingrained Artifact	The overall theory related to ML and DL and their successful implementations in business created the foundation for the project.	
Stage 2: BIE		
Principle 3: Reciprocal Shaping	The team reciprocally solved problems and created development ideas. The data selected was mutually accepted. Members of the team could suggest development ideas.	***Alpha Version***: Initial version based on NN's created. ***Beta Version***: A modified second version of the NN artifact. No significant performance improvements. ***Gamma Version***: Third cycle with a new artifact utilizing UL algorithm SOM provided encouraging results.
Principle 4: Mutually Influential Roles	The ADR team consisted of researchers and practitioners to include theoretical, technical, and practical viewpoints. The researcher and artifact creator worked as an employee in the case organization.	
Principle 5: Authentic and Concurrent Evaluation	Alpha & beta were evaluated with the project team but not released to a broader user setting. Development stopped with gamma.	
Stage 3: Reflection and Learning		
Principle 6: Guided Emergence	Authentic performance evaluation resulted in improvements (alpha to beta) and a new approach to the problem (gamma).	***Current Version and Realization***: A more comprehensive approach to solving AI problems is required. Further development is expected to arrive at testing in a wider setting within the organization.
Stage 4: Formalization of Learning		
Principle 7: Generalized Outcomes	Preliminary design principles were created.	***Ensemble Version***: Artifact and its creation are affected by the team through its observations and organizational practices, legal and ethical considerations, limitations, and restrictions.

6 Discussion

This chapter begins with a conclusion and recommendations attained while creating the artifacts and their evaluation. The chapter proceeds to research evaluation and restrictions and ends with suggestions for future research.

6.1. Conclusions and recommendations

Research began as the needs of the case organization and the researcher's interests met. The case organization expressed interest in AI applications and their usage potential in taxation. The research focused on developing detection models built on AI, such as the study by Chen et al. (2011) and Xiangyu et al. (2018), which provided an ideological starting ground. Consequently, the researcher presented the idea of utilizing AI as an analysis tool in the LLC's claim for adjustment procedure, which was a bottleneck procedure.

This research approached the subject of AI utilization in the Finnish Tax Administration from a holistic point of view and provided answers to the two identified research questions by implementing the ADR method. As a result, two artifacts and a set of design principles emerged. The created artifacts functioned well enough to be considered helpful as knowledge-creating instruments. Therefore, they can create value as knowledge, time, speed, and liberated resources. The instruments analyze data faster than a human being, allowing humans to focus on more urgent matters. AI-based artifacts are investigative objects in the case organization and its activities. However, the challenges in legal and ethical boundaries associated with AI in taxation need to be addressed when creating and using the artifacts. Challenges can be divided into three levels that also form the design principles derived from the ADR method: (1) trustworthiness through accuracy, (2) consideration of legal and ethical restrictions and limitations of use, and (3) justification of use. AI solution achieves trust only by reaching or exceeding a suitable performance. A case organization decides on suitable performance. Legal and ethical

restrictions need to be considered in the whole lifecycle of the artifacts, among other main components. The ADR project with the Finnish Tax Administration required legal and ethical restrictions to be considered while creating and using the artifacts. Different regulative demands most likely occur in other organizations, but some limitations and restrictive boundaries always exist. Lastly, the AI usage should be clearly and thoroughly justified. Therefore a proper preliminary inspection in the problem area considering restrictions and limitations is required. Justifying usage is a primary concern since a usable AI solution cannot be built without a clear view of the problem context.

Due to a holistic viewpoint, a better perception of the problem's context was reached while creating the artifacts with the case organization. According to the research method, even unexpected signals requiring drastic changes to the artifact were followed, resulting in an early version of a different kind of artifact instead of the original one. The emerged design principles emphasize the requirements to be filled to create useful and value-creating artifacts. In the case organization's context, AI usage was limited to knowledge creation with substantial human supervision requirements due to limitations. Based on the results of this research, it is strongly recommended to test out different AI solutions in organizations to understand whether they can offer value of any kind.

6.2. Research evaluation and restrictions

Completely finalized artifacts were not created, and the artifacts were not tested in the end-user setting due to time limitations and the project's status. Created artifacts that worked on top of AI libraries that were intuitively selected. In order to find more suitable libraries, a more profound comparison between different coding libraries would be required. The research process concluded after the third cycle (gamma), and preliminary design principles were created based on the results attained until that point. Moreover, the research would have benefited from testing scenarios in end-user settings and further development cycles for the artifacts to gain more organization-specific information. Consequently, this study should be validated by applying the concept in a case

organization and analyzing the outcomes in an actual work setting instead of a more simulated and restricted one. Furthermore, outsourcing the creation and testing of artifacts would allow more focused roles within the project team.

6.3. Suggestions for future research

Suggestions for future studies include but are not limited to profound research on AI implementation possibilities in taxation, securement of ethical principles in taxation while utilizing AI, and legal perspective on AI usage in taxation. In addition, a comparison of how different AI solutions could enhance automated taxation in Finland could act as a starting point for a future study.

References

Akerkar, R. (2019). Artificial intelligence for business. *Cham: Springer.* https://doi.org/10.1007/978-3-319-97436-1

Alpaydin, E. (2014). Introduction to machine learning (Third edition.). *Cambridge, Massachusetts: The MIT Press.*

Alter, S. (2015). The concept of 'IT artifact' has outlived its usefulness and should be retired now. *Information systems journal (Oxford, England)*, 25(1), 47-60. https://doi.org/10.1111/isj.12048

Benfer, R. A., Benfer, R. A., Brent, E. E. & Furbee-Losee, L. (1991). Expert Systems. SAGE Publications, Inc. https://dx.doi.org/10.4135/9781412984225

Campesato, Oswald. Artificial Intelligence, Machine Learning, and Deep Learning, Mercury Learning & Information, 2020. *ProQuest Ebook Central.* Retrieved 1.10.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?docID=6032875.

Chebbi, Chiheb. Mastering Machine Learning for Penetration Testing: Develop an Extensive Skill Set to Break Self-Learning Systems Using Python, Packt Publishing, Limited, 2018. *ProQuest Ebook Central*. Retrieved 5.9.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?docID=5439458.

Chen, J., Su, M., Chen, C., Hsu, F. & Wu, C. (2011). *Application of neural networks for detecting erroneous tax reports from construction companies*. Automation in construction, 20(7), 935-939. https://doi.org/10.1016/j.autcon.2011.03.011

D'Rosario, M. & Hsieh, C. (2018). Predicting Credit Rating Migration Employing Neural Network Models. *International Journal of Strategic Decision Sciences (IJSDS), 9*(4), 70-85. https://doi.org/10.4018/IJSDS.2018100105

Dutta, Sayon. Reinforcement Learning with TensorFlow: A Beginner's Guide to Designing Self-Learning Systems with TensorFlow and OpenAI Gym, Packt Publishing, Limited, 2018. *ProQuest Ebook Central*. Retrieved 7.5.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?docID=5371683.

Finlay, S. (2018). Artificial intelligence and machine learning for business: A no-nonsense guide to data-driven technologies (Third edition*.). *Retrieved 25.1.2021 from https://www.researchgate.net/publication/326200736_Artificial_Intelligence_and_Machine_Learning_for_Business_A_No-Nonsense_Guide_to_Data_Driven_Technologies_Third_Edition

Finnish Parliament Ombudsman. (2019). Verohallinnon automatisoitu päätöksentekomenettely ei täytä perustuslain vaatimuksia. Retrieved 19.11.2021 from https://www.oikeusasiamies.fi/fi_FI/-/verohallinnon-automatisoitu-paatoksentekomenettely-ei-tayta-perustuslain-vaatimuksia#

Finnish Patent and Registration Office. (2014). Closing a limited liability company. Retrieved 31.3.2022 from https://www.prh.fi/en/kaupparekisteri/lopettaminen/osakeyhtio.html

Finnish Tax Administration. (2015). *Principles of tax assessment procedures*. Retrieved 19.1.2022 from https://www.vero.fi/syventavat-vero-ohjeet/ohje-hakusivu/49044/verotuksen_toimittamisen_periaatteit3/

Finnish Tax Administration. (2016). Tax administration's decision on reporting duties and notes (A123/200/2016). Retrieved 14.12.2021 from https://www.vero.fi/syventavat-vero-ohjeet/paatokset/47259/verohallinnon_paatos_ilmoittamisvelvoll10/

Finnish Tax Administration. (2019a). *Annual report 2019*. Retrieved 11.3.2021 from https://www.vero.fi/en/About-us/finnish-tax-administration/year-2020/earlier-annual-reports/annual-report-2019/

Finnish Tax Administration. (2019b). Apulaisoikeusasiamies: automatisoidun päätöksenteon sääntelytarpeet tulee selvittää viipymättä. Retrieved 6.7.2021 from https://www.vero.fi/tietoa-

verohallinnosta/uutishuone/uutiset/uutiset/2019/apulaisoikeusasiamies-auto-matisoidun-p%C3%A4%C3%A4t%C3%B6ksenteon-s%C3%A4%C3%A4ntelytar-peet-tulee-selvitt%C3%A4%C3%A4-viipym%C3%A4tt%C3%A4/

Finnish Tax Administration. (2019c). *Ethical principles of AI*. Retrieved 20.2.2021 from https://www.vero.fi/en/About-us/finnish-tax-administration/operations/re-sponsibility/finnish-tax-administrations-ethical-principles-for-ai/

Finnish Tax Administration. (2020). Automaattinen päätöksenteko verotuksessa.   Re-trieved 26.3.2021 from https://www.vero.fi/tietoa-verohallinnosta/verohallin-non_esittely/tietosuoja-ja-julkisuus/tietojen_kasittely_ja_tietosuoj/automaatti-nen-paatoksenteko/

Finnish Tax Administration. (2021a). *Processing times for returns and applications.* Re-trieved 26.3.2021 from https://www.vero.fi/en/About-us/contact-us/processing-times-for-returns-and-applications/

Finnish Tax Administration. (2021b). *The largest development project in history com-pleted on schedule and budget.* Retrieved 26.3.2021 from https://www.vero.fi/en/About-us/finnish-tax-administration/year-2020/ensur-ing-tax-revenue/new-technologies-open-up-opportunities/

Finnish Tax Administration. (2021c). *6B Tax return of business activities Corporation*. Re-trieved 26.3.2021 from https://www.vero.fi/en/About-us/contact-us/forms/de-scriptions/6b_tax_return_of_business_activities_co/

Flach, P. (2012). Machine learning: The art and science of algorithms that make sense of data. Cambridge: Cambridge University Press.

Giussani, Andrea. Applied Machine Learning with Python, EGEA Spa - Bocconi Univer-sity Press, 2020. *ProQuest Ebook Central*. Retrieved 26.8.2021 from https://ebook-central-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?do-cID=6110795.

Graupe, Daniel. Principles Of Artificial Neural Networks (3rd Edition), World Scientific Publishing Company, 2013. *ProQuest Ebook Central*. Retrieved 7.8.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.ac-tion?docID=1336559.

Hallintolaki 6.6.2003/464. Retrieved 1.2.2022 from
https://www.finlex.fi/fi/laki/ajantasa/2003/20030434

Jung, A. (2018). Machine learning: Basic principles.
https://doi.org/10.48550/arXiv.1805.05052

Kelleher, J. D. & Tierney, B. (2018). Data science. Cambridge, Massachusetts: The MIT
Press.

Keras. (2021a). About Keras. Retrieved 15.10.2021 from https://keras.io/about/

Keras. (2021b). The Sequential mode. Retrieved 5.5.2021 from
https://keras.io/guides/sequential_model/

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via ma-
chine-learning algorithms. *Journal of Banking & Finance, 34*(11), 2767-2787.
https://doi-org.proxy.uwasa.fi/10.1016/j.jbankfin.2010.06.001

Kirjanpitolaki 30.12.1997/1336. Retrieved 30.3.2022 from
https://www.finlex.fi/fi/laki/ajantasa/1997/19971336

Kohonen, T. (2006). Self-organizing neural projections. *Neural networks, 19(6), 723-
733.* https://doi.org/10.1016/j.neunet.2006.05.001

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural networks, 37*, 52-65.
https://doi.org/10.1016/j.neunet.2012.09.018

Laki elinkeinotulon verottamisesta 24.6.1968/360. Retrieved 30.3.2022 from
https://www.finlex.fi/fi/laki/ajantasa/1968/19680360

Laki varojen arvostamisesta verotuksessa 22.12.2005/1142. Retrieved 30.3.2022 from
https://www.finlex.fi/fi/laki/ajantasa/2005/20051142

Laki Verohallinnosta 11.6.2010/503. Finlex. Retrieved 24.2.2021 from
https://www.finlex.fi/fi/laki/ajantasa/2010/20100503

Laki verotusmenettelystä 18.12.1995/1558. Retrieved 30.3.2022 from
https://www.finlex.fi/fi/laki/ajantasa/1995/19951558

Lee, Wei-Meng. (2019). Python Machine Learning. *ProQuest Ebook Central.* Retrieved 19.1.2022 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?docID=5747364

Maatilatalouden tuloverolaki 15.12.1967/543. Retrieved 30.3.2022 from https://www.finlex.fi/fi/laki/ajantasa/1967/19670543

Minisom. (2022). Minisom 2.3.0. Retrieved 15.10.2021 from https://pypi.org/project/MiniSom/#files

Mirjalili, V. & Raschka, S (2019). *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow 2, 3rd Edition*. Packt Publishing. Retrieved 20.10.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?docID=6005547

Osakeyhtiölaki 21.7.2006/624. Retrieved 1.3.2022 from https://www.finlex.fi/fi/laki/ajantasa/2006/20060624

Pérez López, C., Delgado Rodríguez, M., & de Lucas Santos, S. (2019). Tax Fraud Detection through Neural Networks: An Application Using a Sample of Personal Income Taxpayers. *Future Internet*, *11*(4), 86. MDPI AG. http://dx.doi.org/10.3390/fi11040086

Rahimikia, E., Mohammadi, S., Rahmani, T. & Ghazanfari, M. (2017). Detecting corporate tax evasion using a hybrid intelligent system: A case study of Iran. *International journal of accounting information systems, 25*, 1-17. https://doi.org/10.1016/j.accinf.2016.12.002

Rokach, L. (2009) Pattern Classification Using Ensemble Methods. *World Scientific Publishing Company. ProQuest Ebook Central*. Retrieved 4.5.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia-ebooks/detail.action?docID=1679487.

Sein, Maung & Henfridsson, Ola & Purao, Sandeep & Rossi, Matti & Lindgren, Rikard. (2011). Action Design Research. MIS Quarterly. 35. 37-56. 10.2307/23043488.

Shi, Z. (2011). Advanced artificial intelligence. Retrieved 10.9.2021 from https://ebookcentral-proquest-com.proxy.uwasa.fi

Spyder. (2021). Overview of Spyder. Retrieved 20.2.2021 from https://www.spyder-ide.org/

Statistics Finland. (2021). *Standard Industrial Classification TOL 2008.* Retrieved 1.2.2021 from https://www.stat.fi/en/luokitukset/toimiala/toimiala_1_20080101/?code=A&name=Agriculture,%20forestry%20and%20fishing

Taiwo, O. A. (2010). Introduction to Machine Learning, New Advances in Machine Learning, Yagang Zhang, IntechOpen, DOI: 10.5772/9394. Retrieved 20.2.2021 from https://www.intechopen.com/books/new-advances-in-machine-learning/introduction-to-machine-learning

Tiainen, T., Aittoniemi, J., Haukijärvi, I., & Yli-Karhu, T. (2015). Toimintatutkimus tietojenkäsittelytieteen tutkimuksessa. (Informaatiotieteiden yksikön raportteja; No. 38/2015). http://urn.fi/URN:ISBN:978-951-44-9814-5

Tuloverolaki 30.12.1992/1535. Finlex. Retrieved 24.2.2021. Retrieved 15.3.2022 from https://www.finlex.fi/fi/laki/ajantasa/1992/19921535

Vettigli, G. (2018). MiniSom: minimalistic and NumPy-based implementation of the Self Organizing Map. *GitHub*. Retrieved 15.12.2021 from https://github.com/JustGlowing/minisom/

Xiangyu, X., Youlin, Y. & Qicheng, X. (2018). Intelligent Identification of Corporate Tax Evasion Based on LM Neural Network. https://doi.org/10.23919/ChiCC.2018.8483951

Zhang, L., Nan, X., Huang, E. & Liu, S. (2020). Detecting Transaction-based Tax Evasion Activities on Social Media Platforms Using Multi-modal Deep Neural Networks. . https://doi.org/10.1145/1122445.1122456

Appendices

Appendix 1. Alpha, code sample

```
#Importing libraries
import pandas as pd
import numpy as np
import tensorflow as tf


# Importing data
# Data pre-processing
Y = Tax year has been estimated or not (1=yes, 0=no)
X = All the other values.


# Splitting dataset into training and test set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, shuffle=False, ran-
dom_state = 32)


# Feature scaling non-binary values
from sklearn.preprocessing import QuantileTransformer
sc_X = QuantileTransformer(output_distribution='normal', n_quantiles=5000)
X_train[:,0:5] = sc_X.fit_transform(X_train[:,0:5])
X_test[:,0:5] = sc_X.transform(X_test[:,0:5])
X18[:,0:5] = sc_X.transform(X18[:,0:5])
X19[:,0:5] = sc_X.transform(X19[:,0:5])


# Initializing the ANN
ann = tf.keras.models.Sequential()


# Adding the layers
ann.add(tf.keras.layers.Dense(units=80, activation='relu'))
```

```python
# Adding the output layer. Units = 1 bc binary dependent variable: 0 or 1.
ann.add(tf.keras.layers.Dense(units=1, activation='sigmoid'))


# Compiling and training the ANN
optimizer = tf.keras.optimizers.Adam
opt = optimizer(learning_rate=0.005)
loss_fn = 'binary_crossentropy'
ann.compile(optimizer = opt, loss = loss_fn, metrics = ['accuracy'])


# Adding the output layer. Units = 1 bc binary dependent variable: 0 or 1.
ann.add(tf.keras.layers.Dense(units=1, activation='sigmoid'))


# Training ANN on the training set
ann.fit(X, y, batch_size = 100, class_weight=class_weights, epochs = 10)


# Predictions and model evaluation for 2017, 2018 and 2019
y_pred = ann.predict(X_test)
y_pred18 = ann.predict(X18)
y_pred19 = ann.predict(X19)


# Confusion matrix, 2017 test set results
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
cm = confusion_matrix(y_test, y_pred)
disp = ConfusionMatrixDisplay(confusion_matrix=cm)
disp.plot()


# Precision, recall, f1-score, test set
from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred))
```

```
# 2018 set results

y_pred18 = ann.predict(X18)

y_pred18 = (y_pred18 > 0.5)

cm18 = confusion_matrix(y18, y_pred18)

disp18 = ConfusionMatrixDisplay(confusion_matrix=cm18)

disp18.plot()

print(classification_report(y18, y_pred18))


# 2019 set results

y_pred19 = ann.predict(X19)

y_pred19 = (y_pred19 > 0.5)

cm19 = confusion_matrix(y19, y_pred19)

disp19 = ConfusionMatrixDisplay(confusion_matrix=cm19)

disp19.plot()

print(classification_report(y19, y_pred19))
```

Appendix 2. Beta, code sample

```
# Importing libraries
import pandas as pd
import numpy as np
import tensorflow as tf


# Importing the data
# Data preprocessing


# Splitting data into training set and test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)
```

```python
# Feature scaling 2017, 2018 & 2019 data
from sklearn.preprocessing import QuantileTransformer
sc_X = QuantileTransformer(output_distribution='normal', n_quantiles=1000)

X_train[:,0:5] = sc_X.fit_transform(X_train[:,0:5])
X_test[:,0:5] = sc_X.transform(X_test[:,0:5])
X18[:,0:5] = sc_X.transform(X18[:,0:5])
X19[:,0:5] = sc_X.transform(X19[:,0:5])

# Initializing the ANN
ann = tf.keras.models.Sequential()

# Adding the layers
ann.add(tf.keras.layers.Dense(units=6, input_shape=(3,), activation='sigmoid'))

# Adding the output layer. Units = 1 bc binary dependent variable: 0 or 1.
ann.add(tf.keras.layers.Dense(units=1, activation='sigmoid'))

# Compiling the ANN
optimizer = tf.keras.optimizers.Adam
opt = optimizer(learning_rate=0.05)
ann.compile(optimizer = opt, loss = 'binary_crossentropy', metrics = ['accuracy'])

# Training the ANN on the Training set
ann.fit(X_train, y_train, batch_size = 100, class_weight=class_weights, epochs = 10)

# Predictions and model evaluation for 2017, 2018 and 2019 data
y_pred = ann.predict(X_test)
y_pred18 = ann.predict(X18)
```

```
y_pred19 = ann.predict(X19)


# Confusion matrix, 2017 test set results

from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay

cm = confusion_matrix(y_test, y_pred)

disp = ConfusionMatrixDisplay(confusion_matrix=cm)

disp.plot()


# Precision, recall, f1-score, test set

from sklearn.metrics import classification_report

print(classification_report(y_test, y_pred))


# 2018 set results

y_pred18 = ann.predict(X18)

y_pred18 = (y_pred18 > 0.5)

cm18 = confusion_matrix(y18, y_pred18)

disp18 = ConfusionMatrixDisplay(confusion_matrix=cm18)

print(classification_report(y18, y_pred18))

disp18.plot()


# 2019 set results

y_pred19 = ann.predict(X19)

y_pred19 = (y_pred19 > 0.5)

cm19 = confusion_matrix(y19, y_pred19)

disp19 = ConfusionMatrixDisplay(confusion_matrix=cm19)

print(classification_report(y19, y_pred19))

disp19.plot()
```

Appendix 3. Gamma, code sample

```
# Importing libraries
```

```python
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import minisom


# Data pre-processing as in alpha and beta
# Creating the SOM object
from minisom import MiniSom
som = MiniSom(x=15, y=15, input_len=3, sigma=1, learning_rate=0.5)
som.random_weights_init(X)


# Train the SOM
som.train_random(data = X, num_iteration = 10)


# Visualizing the results (creating the SOM map)
```