T.C.

ALTINBAS UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

IMPROVING OPERATIONAL EFFICIENCY OF GOVERNMENT USING ARTIFICIAL INTELLIGENCE

Saif Salam Ibrahim AL-WITWIT

Master of Science

Supervisor

Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Istanbul, 2020

IMPROVING OPERATIONAL EFFICIENCY OF GOVERNMENT USING ARTIFICIAL INTELLIGENCE

by

Saif Salam Ibrahim AL-WITWIT

Electrical and Computer Engineering

Submitted to the Institute of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2020

The thesis titled **"IMPROVING OPERATIONAL EFFICIENCY OF GOVERNMENT USING ARTIFICIAL INTELLIGENCE"** prepared and presented by **Saif Salam Ibrahim AL-WITWIT** was accepted as a Master of Science Thesis in Information Technologies.

Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Supervisor

Thesis Defense Jury Members:

Asst. Prof. Dr. Abdullahi Abdu IBRAHIM	School of Engineering and Natural Sciences, Altinbas University	_____
Prof. Dr. Osman Nuri UCAN	School of Engineering and Natural Sciences, Altinbas University	_____
Asst. Prof. Dr. İzzet Paruğ DURU	Vocational School , Istanbul Gedik University	_____

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Approval Date of Institute of Graduate Studies:
____/____/____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Saif Salam Ibrahim AL-WITWIT

DEDICATION

I devote and pledge this research work to my supervisor who is salient for guiding me through whole research work as well as my family for always assisting me in my hard time.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Asst. Prof. Dr. Abdullahi Abdu IBRAHIM for guiding and helping me along the way in writing this dissertation. Discussing my progress, problems, and ideas with my supervisor Asst. Prof. Dr. Abdullahi Abdu IBRAHIM a couple of times every week helped me tremendously in understanding the logic behind the research. It made me better realize the technical need for this research work. He helped me by looking outside of all the machine learning and network-based information, and considering the cryptocurrency-based knowledge from a more holistic perspective than just from the productivity point of view. He patiently explained to me a great number of technical details with regards to block-chain and cryptocurrency, for example how to deal with solving problems and to provide insight into the reliability of those results. Without his time and help none of this research would have been possible, and I'm grateful to him for giving me the opportunity to work on this interesting subject. I am also very thankful to my parents as well as my friends who always supported me through whole of this journey.

ABSTRACT

IMPROVING OPERATIONAL EFFICIENCY OF GOVERNMENT USING ARTIFICIAL INTELLIGENCE

AL-WITWIT, Saif Salam Ibrahim,

M.Sc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Date: November /2020

Pages: 61

In this thesis we developed a technique for Improving the operational efficiency of the government using Artificial Intelligence, we discussed the concept of insurance and how the concept of insurance will change and Artificial Intelligence (AI) is already disrupting the state of this industry. Insurers worldwide are using AI to automatize processes and tasks, such as fraud detection, underwriting, and claims processing. Additionally, there has been a rise of new competitors in the market, such as InsurTechs, that are bringing innovative solutions for insurance using Artificial neural network (ANN), responding to the new trends in customers' lifestyles and behaviors, that are more demanding for services directed for their needs. This study aims to understand how personalization of insurance policies, created with Artificial Intelligence and how its efficiency can be improved, and how it will disrupt this industry in the future and what will be the impact on the government's operational efficiency. We have chosen worldwide Governance indicator dataset which is publically available the personalization of an insurance policy with AI would encompass the definition of the coverages and premiums more appropriate for an individual customer and do the risk evaluation, in a market of one strategy. This innovation would take advantage of the accrual of Big Data from customers for the optimization, as people are each time more connected and information about them is constantly being shared, allowing companies to use it to know

consumers better and for the training, testing and validation a well-known MATLAB R2019a software was used for this purpose. We achieved an accuracy of 95.25% using 9-Fold cross-validation.

Keywords: Artificial Intelligence, ANN, Big Data, optimization.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

LIST OF ABBREVIATIONS

AI	:	Artificial Intelligence
BIS	:	Bank of International Settlement
CPP	:	Cryptocurrency Price Prediction
DLT	:	Distributed Ledger Technology
FTR	:	Funds Transfer Regulation
FIR	:	Fixed Investment Rate
LCL	:	Legal Cryptocurrency Law
ML	:	Machine Learning
OTF	:	Organized Trading Facility
PoS	:	Proof of Stake
PoW	:	Proof of Work
TCI	:	Technical Trade Indicator

1. INTRODUCTION

1.1 INTRODUCTION

Industries all over the world are being disrupted by technological advances that are transforming their existence, and the way companies react and adapt to these disruptors will have massive impacts on how industries will look like in the future.

An example is the Insurance industry that has been under the disruption of new trends, such as Artificial Intelligence (AI), driven by the quick advances in the technological field. This transformation includes several other disruptive technologies, such as telematics, Internet of Things, blockchain and digital platforms and the accumulation of Big Data has been feeding these advances.

Technology based innovations have several benefits for insurers, for example, for controlling risks, cost efficiency, better engagement with its customers and tailoring insurance offers [21]. With the evolution of technologies, insurance will change its current state from a “detect and repair” to a “predict and prevent” approach.Some of the main issues to tackle in insurance are technology advancements, regulatory implementations, product development, mergers and acquisitions, privacy issues and tax reforms, which insurers must take into consideration in order to improve operational efficiency, increase productivity, lower costs and customize their policies .Customers’ expectations towards companies and the desire for personalized and flexible insurance policies is emerging and insurers must assess their ability to respond to it. As a result, an alternative for insurers is to join forces with InsurTechs to implement new approaches, platforms and policy plans.

Every day, millions of data about customers’ lifestyle and behaviors are generated and accumulated as Big Data. Artificial Intelligence has been a solution to transform and analyze the high amounts of information that can be further used to create personalized policies based on client’s exact need, for different categories.The creation of this innovation requires high amounts of data relative to customers, which can originate privacy constraints, in which individuals may be withdrawn to share. Furthermore, as the insurance industry is highly regulated, Turkey insurers must consider the regulation applied for their activity and regarding discrimination and data

protection. In Turkey, an example is the General Data Protection Regulation (GDPR), created by the Turkey Government Personalization of insurance policies using Artificial Intelligence may be an opportunity to increase the number of policyholders by increasing the intention to acquire an insurance, as many people still do not want to acquire it, whereas because they consider themselves to be healthy – in the case of health Insurances, or due to its high prices or unnecessary coverages. In addition, it can lead to increases in revenues and accelerate industry growth.From the managerial relevance of this study, it is pertinent to highlight the Insurance industry and the AI developments. Insurers contribute to the economic growth, financial stability and development of Turky and companies search each time more for solutions that embed AI benefits in their strategies, especially for reducing costs and increasing efficiency.The academic relevance of this study is to present a recent and in high demand topic to the academic community. In addition, it intends to deliver insights on how to define strategies to overcome or leverage disruption in industries. Life Insurance is a contract in which the policyholder acquires an insurance that will be reimbursed upon death or maturity, being considered a form of investment. It is a complex service, often associated with uncertainty of its unsure future benefits, and the way customers choose it depend on their proposal, agent, image of the insurer company and it also varies according to the culture of the country [8]. General insurance consists in all types of insurance of non-life policies (e.g. health, work, property, vehicle), and consist in reimbursements for uncertain events. Non-Life insurance products are standardized and prices follow actuarial principles. The price definition involves risk evaluation of the insured person or object, that is private to the firms. Insurers hold a portfolio of insurance policies and an investment portfolio, and the income of insurers come from two sources. The first one is from the portfolio of customers insured, originated from the underwriting process, that is the difference between the premiums paid from customers and the payments made to the policyholder for incurred losses and expenses. There exists a concept of mutualism, based on a principle of pooling risks, in which policyholders who do not have losses subsidies those who do. The second source is from insurers investing the premiums earned in common equities and fixed-income securities.

1.2 ARTIFICIAL INTELLIGENCE

American pioneer in the field of Artificial Intelligence, Arthur Samuel, is responsible for introducing the term machine learning in 1959 [1]. Machine learning works on prediction using

computer and data and also associated with computational statistics. ML is a sub-category of AI. ML contains different theories, methods and application domains as it has a strong connection with mathematical optimization.

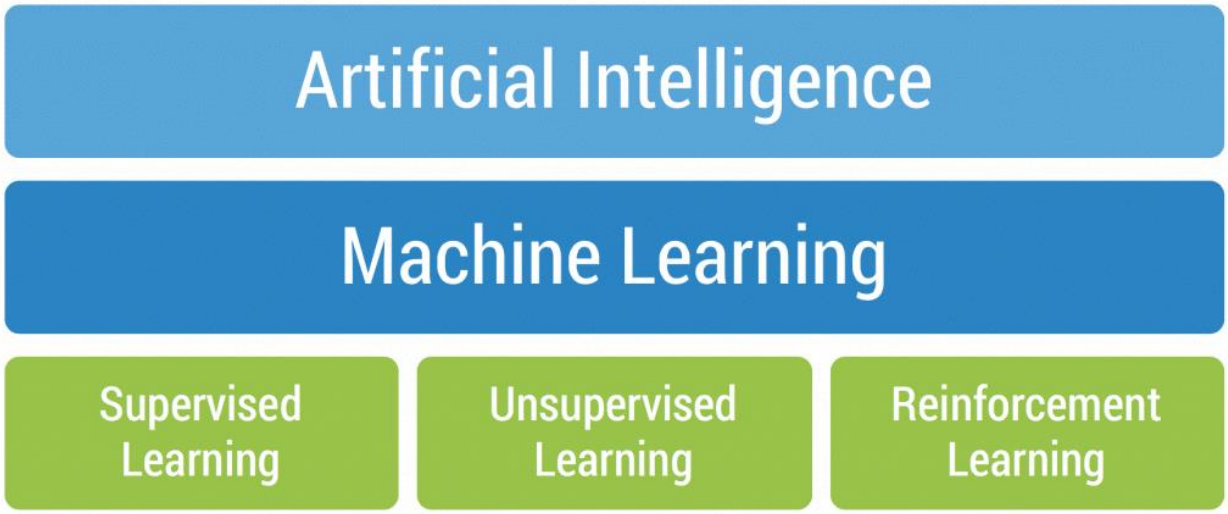


Figure 1.1: Taxonomy of artificial intelligence and machine learning with three approaches of machine learning.

Machine Learning primarily includes the following four steps:

Choose the basis attributes or features or values for prediction. Choose the appropriate machine-learning algorithm according to the worldwide Governance indicator dataset. For example, among classification algorithm or regression algorithm, which one to choose, high complexity or faster one.

- Training and evaluating model performance based on different algorithms.
- Classify or predict the unknown data, training model is used.

1.2.1 Supervised Learning

Algorithms for supervised learning are used to build a model from a set of data that has both the desirable inputs and outputs. The data used is training data and contains training examples that has single or multiple inputs and an output [2]. The method used is known as supervisory signal.

When dealing with problems regarding supervised signal, first we begin with a set of training examples with correct labels associated with it. An example is when we train a supervised learning algorithm which takes thousands of pictures of hand- written digits with the correct labels containing the correct number for each image it represents [3]. This enables the algorithm to classify handwritten digits and also learn the relationship of the images and their respective numbers and also use this relationship to classify new images without labels that has not been shown to the machine before.

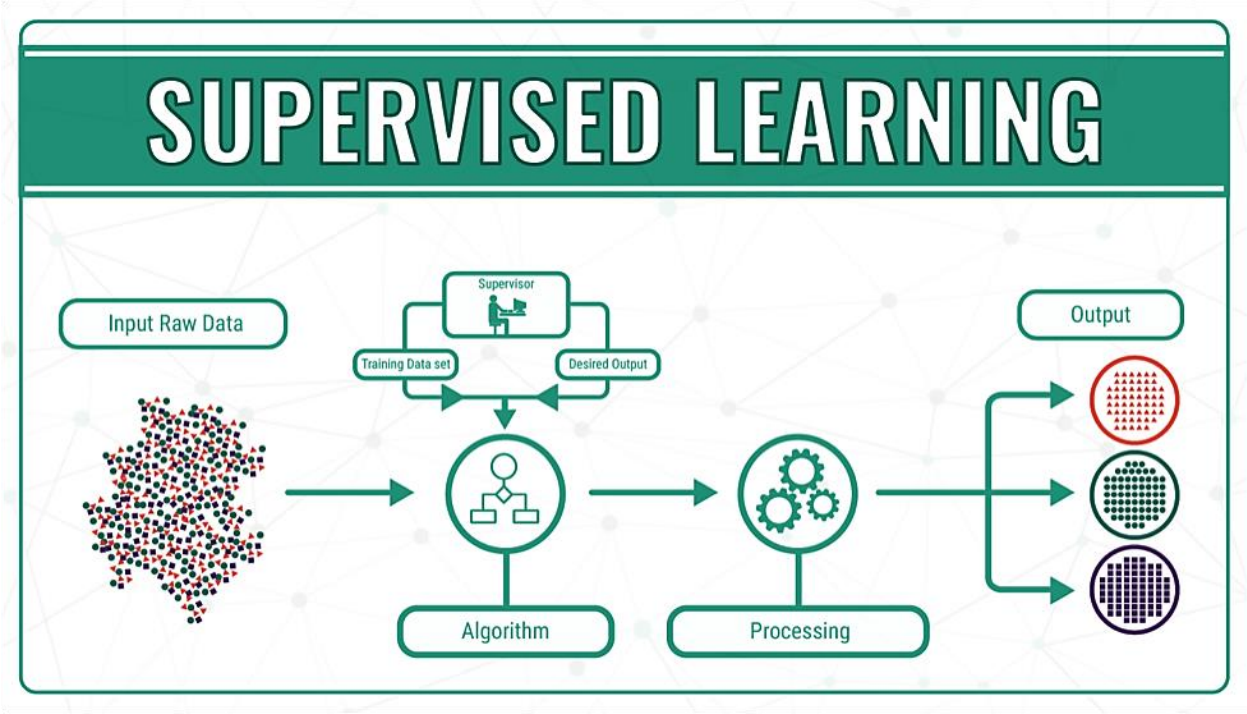


Figure 1.2: The training of supervised based algorithm from input to output. Source [19]

Classification and regression is a part of supervised learning [4]. Algorithms for classifications are used in cases where outputs are restricted to sets of values which are limited. Algorithms for regression are used in cases when outputs may have values which are numerical within a given range.

1.2.2 Unsupervised Learning

Algorithms for unsupervised learning are given sets of data that contain inputs only and then the algorithm finds clustering of data points or finds structure in data [5]. Therefore, the algorithms

learn from non-labeled, non-categorized or non-classified test data. Unsupervised learning algorithms detect commonalities in the data which is based on the absence or presence of the commonalities in each new piece of data instead of a feedback approach.

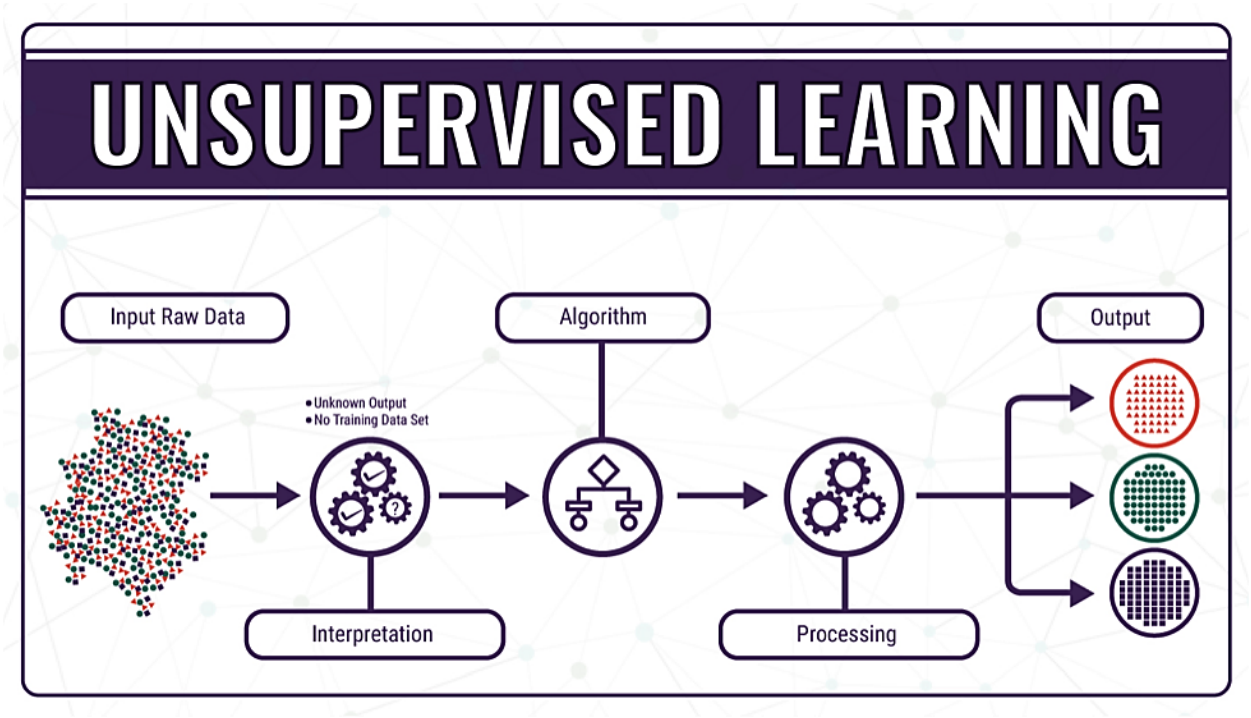


Figure 1.3: The training of unsupervised based algorithm from input to output. Source [19]

Grouping, summarizing and finding underlying structure of unlabeled data in a useful manner is the function of unsupervised learning.

1.3 RESEARCH QUESTIONS

The problem this dissertation attempts to understand is how a market of one strategy, with personalized insurance policies by Artificial Intelligence, will affect the future of the Insurance industry in Turkey. In order to conduct the study and hypothesize the future of this industry, the following research questions will be answered:

RQ1. What is the impact that AI will have in the Insurance industry in Turkey?

RQ2. How is this industry going to look like in the future, once this disruptor has taken full effect?

In order to research the problem and answer the research questions, both primary and secondary methods of data collection were used. The study consists in an exploratory approach and it used qualitative and quantitative data. For the primary data collection, a survey was done with Turkey respondents and semi-structured interviews were conducted with experts working in Insurance and Artificial Intelligence. For the secondary data, the theoretical foundation of the study was supported by existing literature, such as academic articles, journals, companies' reports, industry indicators in Turkey government and other relevant studies [6].

1.4 AIM OF THE THESIS

The purpose of this study is to improve the operational efficiency of Government using Artificial Intelligence. In this thesis we examine the future of the Insurance Industry in Turkey, by identifying the major disruptors in this industry, with focus on Artificial Intelligence. More specifically, it intends to assess how AI powered solutions can be used to personalize insurance policies, in a market of one strategy. Personalization would encompass creating a policy suitable to a customer's individual needs, involving the definition of the most appropriate coverage extent, capital available, premium and evaluation of the risks [7].

1.5 PLANNING OF CHAPTERS

The planning of Chapters will be that Chapter 1 will present an overview of the thesis, as well as a short summary off each chapter. Chapter 2 will present background knowledge and related work that is relevant to the topics of the thesis. The Operational Efficiency of Government is improved in this thesis using Artificial intelligence. Following this, several topics related to operational efficiency of government will be describe Chapter 3 will present a theoretic overview of the method proposed in this thesis. It presents the structure of the method and the proposed framework for implementation and comparison, and suggests several sub-methods that can be used to implement it. An open source worldwide Governance indicator dataset is used and artificial neural network algorithms will be used to train in Matlab and evaluate the proposed method. This dataset will also be presented. Chapter 4 will present the results and lab practice are presented and compared in this chapter. Chapter 5 will present the final chapter presents a summary and conclusion of the thesis. This chapter provides a discussion off the results, as well as the

considerations and implications of the thesis, and a description of future research that may be conducted to further study the subjects presented, and in the end, there will be References.

2. RELATED WORK

2.1 THE INSURANCE INDUSTRY.

The Insurance industry is composed by two main groups of policy contracts: life and non-life (general) insurance. The insurer pays a compensation to the policyholder in case of loss, damage or death, in return of a regular payment of the premium. Insurance policies contain coverage for losses arising from uncontrollable factors and allow risk averse clients, due to wealth constraints, to transfer risk to an insurer that takes a risk neutral role.Life Insurance is a contract in which the policyholder acquires an insurance that will be reimbursed upon death or maturity, being considered a form of investment. It is a complex service, often associated with uncertainty of its unsure future benefits, and the way customers choose it depend on their proposal, agent, image of the insurer company and it also varies according to the culture of the country.General insurance consists in all types of insurance of non-life policies (e.g. health, work, property, vehicle), and consist in reimbursements for uncertain events. Non-Life insurance products are standardized and prices follow actuarial principles. The price definition involves risk evaluation of the insured person or object, that is private to the firms [8]. Insurers hold a portfolio of insurance policies and an investment portfolio, and the income of insurers come from two sources. The first one is from the portfolio of customers insured, originated from the underwriting process, that is the difference between the premiums paid from customers and the payments made to the policyholder for incurred losses and expenses [9]. There exists a concept of mutualism, based on a principle of pooling risks, in which policyholders who do not have losses subsidies those who do. The second source is from insurers investing the premiums earned in common equities and fixed-income securities[10].

In the last decades, the Insurance industry has been under several transformations, due to globalization, deregulation and the massive digitalization (Mitra, 2016). Nevertheless, insurance activities have been growing, especially in emerging markets, driven by the process of financial liberalization and integration that contribute to economic growth, as a result of the premium incomes, insurers' assets and investments made[11].

The Insurance industry is extremely regulated and traditional. Its customers consider it impersonal and with lack of enthusiasm , with low consumer engagement and interaction, since there are only

two points of contacts with customers: when the customer buys the insurance and during claims processes[12].

2.1.1 Turkey Market

The Turkey insurance market is the largest in the world, representing over 30% of the global insurance premium income . Hence, the Turkey market is being studied due to the competitiveness and efficiency of its institutions. This market is composed by large and multinational companies, operating in life and non-life segments, through numerous subsidiaries with different independence levels to conduct their business and that co-exist with other fully independent firms that operate in both segments [13]. Insurance companies typically have growing returns to scale, which originated increases in firms' sizes and market concentration. Additionally, due to the liberalization of the EU market, the number of mergers and acquisitions among companies increased [14].

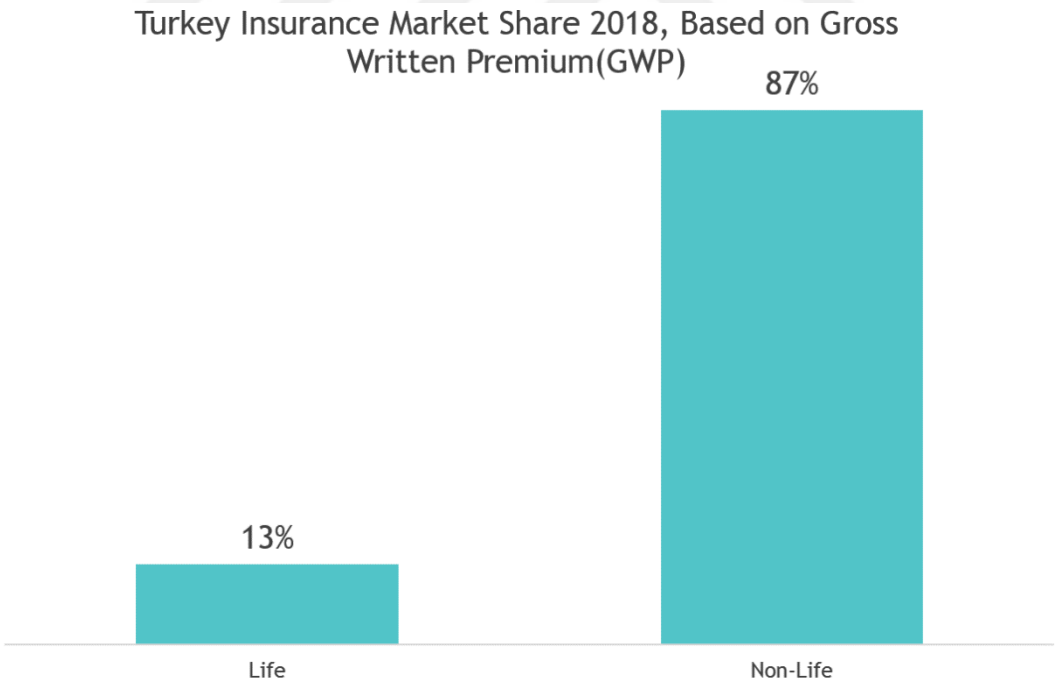


Figure 2.1: Turkey insurance market

The financial services sector, that includes insurance, has been gradually being deregulated by the Turkey with the objective of creating a single market in financial services. This measure allows insurers to do business in every EU country provided with the respective licence [15].

Additionally, the Insurance sector allows countries to perform economic transactions among them (e.g. risk transfer and long-term investments) . The insurance sector is the largest institutional investor in Turkey and provides economic growth (Insurance Turkey, 2019), financial stability and development of the nations .

2.1.2 Factors Influencing the Consumption of Insurance Products

In the case of life insurance in Turkey, the demand is determined by economic, demographic and cultural variables. The economic factors are countries' GDP per capita, and consequently its higher wages and higher activity levels. Better developed banks also increase the consumption of life insurance. The demographic variable that positively impact the demand is the size of the country's population and a cultural variable of long term orientation (Mitra, 2016).

2.2 ARTIFICIAL INTELLIGENCE AS THE MAIN DISRUPTOR

Currently, we are living in the era of Big Data, as a result of the fast advances in technology over the last decades and the diffusion of large amounts of information online[16] New technologies, such as Artificial Intelligence, has been used by diverse industries to develop their businesses and Artificial Intelligence was listed as the number one strategic technology in 2018 .Artificial Intelligence is the capability of a machine to learn from experience and inputs, in order to perform tasks similarly to a human . These tasks include learning, planning, reasoning, problem solving and decision making, and it may support or completely substitute human tasks .

This innovation has been expanding as a result of the improvements in computer's processing and memory capacity, as well as the accumulation, availability and power of Big Data .Artificial Intelligence developments have been enabling to process and run complex algorithms, faster than humans could , as Big Data analysis is extremely time-consuming .This advance has been helpful for companies to discover new insights, shorten processing times, improve service and products quality and use data to make predictions, and as a result, it contributes to reduce the cost of these activities ,Consumers generate each time more data that can be either structured or unstructured. Some examples of sources of data that generated new types of information are social media, wearable devices, sensors and telematics, that permit AI to be useful to process it and generate new ideas based on consumers' information [17].

Firms' innovations in R&D are becoming less science-based, decreasing the benefits of investing in scientific research, as it is not considered as relevant as it was years ago for commercial purposes [18]. Artificial Intelligence and other IT advances are a result of statistics, computer science, electronical engineering and material science, and its connection to science is still important to take into consideration, as these types of technologies are rooted in advances made from science.

2.2.1 Artificial Intelligence Techniques

The core AI components are Machine Learning and Deep Learning. Following, there are AI enabled technologies, such as Natural Language Processing, Machine Vision and Predictive Analytics [19].

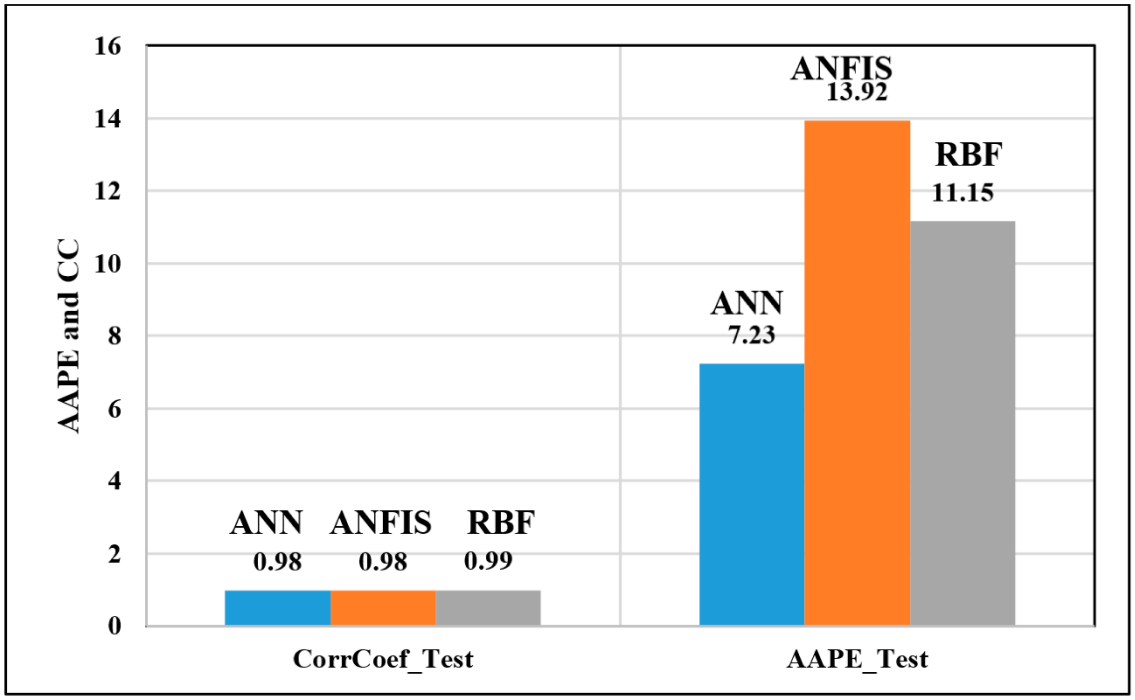


Figure 2.2: Artificial Intelligence Applications

In addition, companies do Data Mining to process Big Data.Some of the main techniques used in AI's systems to develop its knowledge and algorithms are rule-based inference, semantic linguistic analysis, Bayesian networks, similarity measures, neural networks, frame-based representation and genetic algorithms [20].

2.2.2 Future Predictions for AI

Artificial Intelligence is self-learning and over the time these systems adapt to new data and improve their capabilities (Hall, 2017).

Nowadays' Artificial Intelligence is considered as weak, being designed to perform a specific task and being able to outperform humans on that specific task, and innovation technologies for Deep Learning and Corporate Cooperation are emerging[21].Today's Artificial Intelligence developments only encompass the intellectual areas of image and speech recognition and dialogue response of the human brain, which limits its full potential. In addition, it includes several techniques that are only part of the areas of the human brain, and still does not perform like it. For instance, AI's solutions are not yet capable of performing the self-functions of understanding, control, consciousness and motivation that humans do .The recompense for digital initiatives through 2025 will be driven by the ability to use AI for decision making, reinventing business models and ecosystems and for customer experience[22].

2.3 ARTIFICIAL INTELLIGENCE APPLICATIONS IN INSURANCE

Technology is developing in a fast pace and its advances generated new sources of data and opportunities to leverage. By this reason, Artificial Intelligence is disrupting several industries worldwide, an already has outstanding effects in the Insurance industry.

By using algorithms, firms can understand better their consumers' behaviour and characteristics, which is beneficial to create value not only for the firm, but for customers, generating loyalty to the brand. In addition, using data to serve customers is more difficult to imitate from other firms, which benefits the creation of new and unique innovations [23].Nowadays' consumers, especially the Millennials generation, expect on demand services, and AI is being used for improving customers experience and innovating products. For this reason, personal data from customers can be used to increase engagement with insurers and make the relationship more dynamic and interactive, modernize and personalize policies .

2.3.1 The Rise of InsurTechs

Technological disruption is reshaping the financial services sector, which created points of contact between traditional financial companies and start-ups, originating FinTechs (financial technology driven companies). In the specific case of the Insurance industry, it is being disrupted by InsurTechs, a sub-group of FinTechs, that are using innovative technology and disruptive business models in Insurance, making use of Big Data, AI and Internet of Things methods as a way to differentiate products, processes and delivery models[24]. Artificial Intelligence is being used for underwriting, fraud detention, claims processing, marketing strategies , algorithmic trading, blockchain analytics and customer service improvements with chatbots [25].The number of InsurTechs has been increasing and is in high demand for improving customer experience with convenience, transparency, timeliness and engagement . Synergies between InsurTechs and traditional insurers can be leveraged for product development, as InsurTechs can create products with more appropriate coverages, better platforms and without annual commitments. In addition, this type of business model has lower burden of regulation than a traditional insurer. On the other hand, InsurTechs can benefit from insurers' stablished position in the market, for its brand recognition, experience and capital accessibility[26].

2.3.2 AI Techniques used in Insurance

In order to automatize processes and personalize products in Insurance, several AI-powered solutions are used. Starting with underwriting activities, AI techniques are used to asses risk and price the policy in accordance. Typically, insurers use information such as age, gender, health state, among others, but with AI techniques it can use real time information from different sources and have better information about their customers. Sensors can be used for assessing customer's behaviour in relation to the risk being insured and tailoring the insurance policy to the specific customer risks, corresponding to real data [27].

Natural Language Processing is used in Insurance to create chatbots, that are virtual assistants that answer to customers' enquiries through messaging platforms, instead of having a human performing this task and it aims to improve customer service[28].Another technique in high demand is Data Mining that is used for marketing and market analysis to support the high amount of data being generated every day. Each time more,consumers are online and share different types

of information about themselves (e.g. on social media), hence, one of the methods used in Data Mining is through consumers' online searches, browsers' information and products ordered online, as a way to identify their purchasing behaviour pattern. For marketing purposes, it is used for advertisement and for recommending products. On the other hand, Data Mining permits to do analysis of the risk of the customer and detect fraud probability, through customer's creditworthiness or fraud insurance, for instance[29].Claims processing times are being shortened by using Machine Learning techniques, that uses bots to analyse historical data, images and sensors, and permits to automatically review the claim (e.g. evaluate the severity of the claims and predict costs), cross it with the customer's policy, perform anti-fraud algorithms, approve it and send wiring instructions to the banks. At the end, it informs the policyholder that the order was paid and closed. This also decreases the costs associated with it and improves customer experience [30].

Artificial Intelligence also aids to perform anti-fraud techniques to detect doubtful claims, through intelligent automations, self-learning and detecting patterns to identify fraudulent claims. This works by crossing information with the customer's policy, assessing previous customer behaviours and detect activities with high probability of being fraudulent, which sometimes is impossible to detect through humans [31].

2.3.3 Personalization of Insurance Policies using AI

Nowadays, there is an array of devices that are all connected and generate data about consumers, such as cars, home assistants, fitness trackers, smartphones and watches. In the future, there will be even more information being shared and by 2025, it is forecast that there will be one trillion of connected devices and this connectivity will be extended for clothes, shoes, eyewear, medical devices, among others[32] The data being shared accrues as Big Data and is further used by AI for creating products and services, allowing insurance transactions to be personalized through innovations in social media marketing and behaviour tracking, for instance[33].This will allow companies to collect higher amounts of information about consumers and define better strategies based on a deeper understanding of their needs and preferences, which will empower the possibility to innovate in the personalization process of new products.A common segment in which technological innovation is being applied in the car insurance, through the application of

telematics, a technology that combines telecommunications, wireless and vehicular technologies to monitor customer's driving behaviour and determine the premium of their policy accordingly. Historically, this type of policies used actuarial prices and divided drivers into groups, based on historical facts about claims and considering information such as age, gender, driving licence and vehicle type. This led younger drives to pay more due their less driving experience and association with higher claims costs. Alternatively, telematics policies aim to be personalized based on individual data, considering drivers' locations, speed and driving behaviour[34]In addition, sensors are also being used on cars to measure acceleration, braking, time of the day driving and upload all the information on the company.Another example is in health and life insurance that is extracting data about policyholders from wearable devices and health apps to monitor their activities such as steps taken, exercise, diet and sleeping patterns, which allow insurers to calculate risks and modify premiums to charge them. An example is the Vitality program, created by the South African insurer Discovery, that rewards its customers for achieving points for their healthy behaviour and gives policyholders the possibility to decrease their premiums based on their points achieved.The British insurer Aviva created the product Ask It Never, an underwriting innovation that permits customers to buy policies without asking them endless questions. Instead, it relies on Big Data, where digital footprint data takes an important role to underwrite the risk.

Another case is the Selfie-Quote.com, where customers upload a photography of themselves and though facial analytics technology, it indicates an estimated quote for life insurance based on the user's age, gender and Body Mass Index. This project was a collaboration of Legal and General America and a technology firm [35].

2.3.4 Usage-Based Insurance

An innovation that some insurers adopted is the Usage-Based Insurance (UBI) that are personalized to consumers and adapted to their individual behaviours, with coverages based on their needs and annually renewable. This model includes micro coverage, that includes coverages for phone battery or flight delays, for example, based on the customer's needs. Nowadays, with the emergence of the shared economy, this model is highly appreciated for physical assets that are shared, for example, cars and home services. Shared economy has several risks associated and businesses operating in this model prefer to have coverages they can activate and deactivate when

needed[36]. Furthermore, customers want more control over their coverages and to tackle this opportunity,InsurTechs are already creating coverages based on real time needs. An example is the on-demand insurer Trōv that allows individuals to use an insurance for objects and activate the coverage through a mobile app whenever they need (e.g. an insurance for cameras that is activated when the customer is traveling

2.3.5 Limitations of the use of AI in Insurance

Telematics is being used to define the premium for car insurance policies. Telematics also include detection of other behaviours, such as phone usage and alcohol levels. In this sense, it can create issues regarding who takes the responsibility, because this technology may fail or relate false positives that can have negative impacts for the customer and for the image of the insurer [37].Using self-tracking data to analyse and price individual risk has several obstacles. Because the Insurance industry is heavily regulated, it has several legal prohibitions that limits the type of personal data that insurers can use to price risk.Some additional challenges are related to the quality of the data gathered, infrastructure compatibility and privacy constraints as personalization also involves some restrictions associated to the privacy and security of personal data shared.For storing data, the cloud takes a very important role because it can storage large amounts of information and keep it safe. However, companies must bear in mind the risk associated with cybersecurity and insurers must take care of the implication of migrating data to cloud and risk management.

2.4 ARTIFICIAL INTELLIGENCE

Deep learning is a branch of machine learning based on neural networks. It could be supervised, semi-supervised or unsupervised. Supervised learning is the one that learns labelled training data as an example and maps an output based on the level of learning. The training data is analysed, and an inferred function is produced that could be useful for generating newer examples. In ideal terms, a supervised algorithm can accurately determine class labels for unobserved instances. Similarly, semi-supervised learning is a class of machine learning that uses both labelled and unlabelled datasets for training purposes. However, the amounts of unlabelled training datasets are far more compared to the labelled set of training data. It has been observed that the learning accuracy is considerably improved when larger unlabelled set of data is used with smaller set of

labelled data [38]. Likewise, unsupervised learning is a machine learning technique that figures out unidentified outlines of dataset without any pre-existing labels. This type of learning mainly uses the principal component and cluster analysis. For Principal Component Analysis, an orthogonal transformation is used in order to alter a group of observations of correlated variables to linearly uncorrelated variables, also referred to as principal components. Similarly, for cluster analysis, the data that has not been labelled is grouped and is further clustered or categorized to evaluate common ground in datasets. This in fact, helps detect anomalous objects that do not appear familiar or do not fit into the group.Basically, deep neural networks are in fact artificial neural networks with multiple hidden layers in between the input and output layers. Due to the presence of multiple layers, deep neural networks are capable of modelling complex non-linear relationships with complicated datasets which most similarly, performing single-layered shallow networks fail to do. The main advantage of working with deep learning algorithms is that they do not need structured or labelled data to perform prediction[39]. Deep neural networks rely on the hidden layers of artificial neural networks and learns through each layer hierarchically. The neurons of deep learning are classified into 3 different categories:

1. Input Layer
2. Hidden Layer
3. Output Layer

Figure 2.3 shows the graphical representation of deep neural networks. The input data is received through the input layer where it transfers the data to the first hidden layer. In Figure 2.3, we have altogether 4 different hidden layers. All the mathematical computation occurs inside the hidden layers. The data is computed in each first hidden layer and its output serves as the input of the second hidden layer and so on. It is therefore, important to understand what number of hidden layers and the number of neurons in each of these hidden layers would help get the best overall prediction. Similarly, after successive computations in each of the 4 hidden layers as shown in figure 2.3, the output layer returns the final output data, which we call the predicted value.

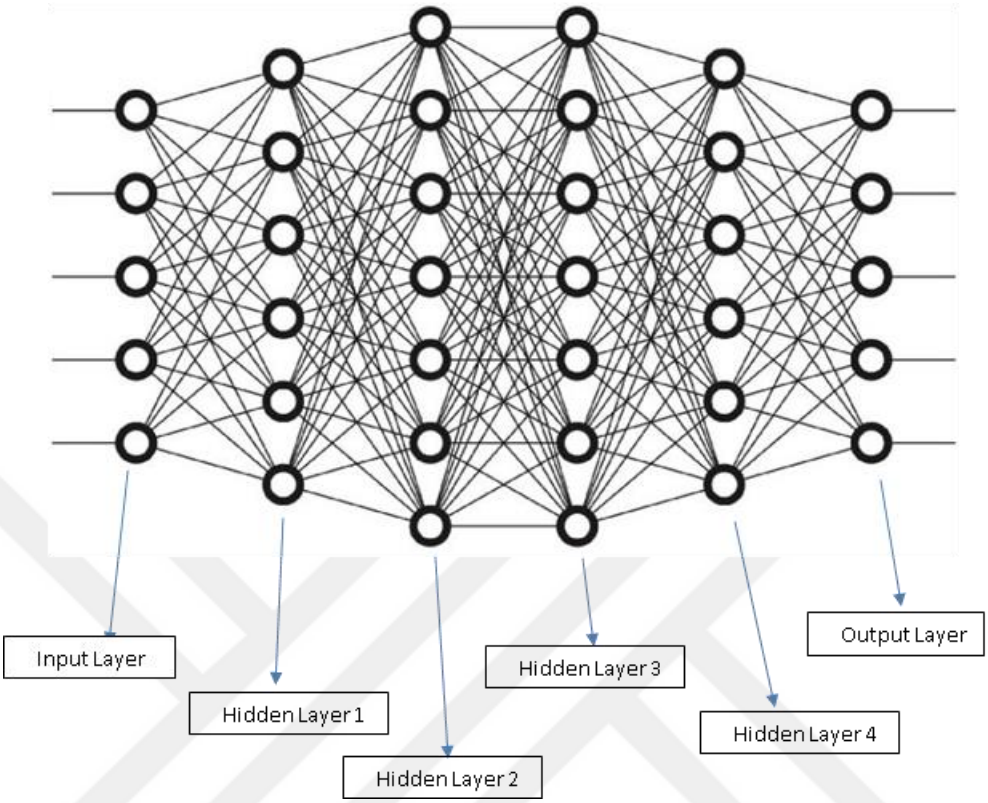


Figure 2.3: Artificial neural network

An image is passed through a Artificial neural network and assigned importance based on learnable weights and slight biases to certain features or items in the image that can differentiate one part of the object from another in same or dissimilar images. The pre processing time required in deep convolutional neural networks is comparatively lower in comparison to other classification algorithms. Convolutional Neural Networks can purposefully seize both spatial and temporal dependencies throughout the image by applying numerous relevant filters. Especially, with the advent of GPU and improved algorithms, deep neural networks usage has surged several folds for analysing visual imagery. Some of the algorithms used in the current study are described as follows.

2.4.1 Artificial Neural Network

ANN uses the processing of a brain to develop algorithms and the methods based on artificial Neural Networks (ANNs) have proven especially effective in Video Object Tracking, With the increase in computational power of modern computers and hardware, and an increase in data

availability, complex models are able to achieve both high levels of accuracy and low processing times. With the correct model and powerful hardware, real-time performance can be achieved. The main goal of this project is to develop a video object tracking system with ANN that is able to run on the raspberry pi system in real-time. The concept of classical programming is that an engineer defines a set of rules, called an algorithm, which uses input data to calculate some form of output data [40]

Figure 2.4: Classical programming pipeline

A machine learning algorithm is an algorithm that can learn from data. It can be used to calculate these rules automatically, so they do not have to be specified by hand. Three components are needed for such an approach. It works by feeding input and output data into a pipeline, which will learn to transform one into the other. With the advantage that no explicit pro- gramming is needed to generate the rules, comes the disadvantage that prior input and output data is required for the initial learning process.

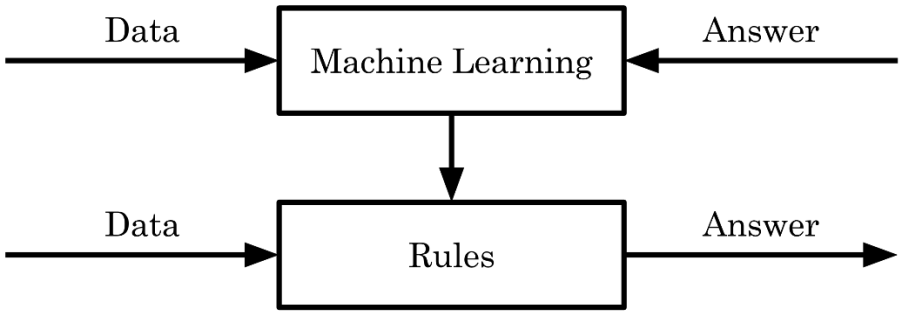


Figure 2.5: Machine learning pipeline

Machine learning may be applied as an effective method if it is not feasible or possible to define an algorithm by hand and sufficient data is available for training. How much “sufficient” is depends on factors like the type of task, the complexity of the data, the uniformity of the data, the type of machine learning algorithm and others.

There are different subparts to machine learning like supervised and unsuper- vised learning. Supervised learning is used when it is clear what the output data looks like, whereas unsupervised

learning can help to find unknown patterns in the data. Examples of supervised learning techniques include lin- ear regression, naive Bayes, support vector machines, decision trees, random forests, gradient boosting and artificial neural networks (ANNs). Since the primary interest of this study revolves around ANNs, this will be the focus of following chapters.

3. METHODOLOGY

In this research work, Collection of raw data is very hard to do. Moreover, working with raw data is tough as it contains many repetitive rows and anomalous values which do not reflect the true situation. Therefore, raw data has to be filtered and molded into worldwide Governance indicator dataset that can be utilized according to the researcher's purpose. Similarly, our data is filtered to remove rows that are exactly same. Synthetic values are introduced in order to reflect better results on the sets. The purpose of this thesis is to developed a technique for Improving operational efficiency of government using Artificial Intelligence, how personalization of Insurance policies, created by Artificial Intelligence, will affect the Insurance industry status-quo. The research questions that the study aims to address are: RQ1. What is the impact that AI will have in the Insurance industry on government operational efficiency?

RQ2. How is this industry going to look like in the future, once this disruptor has taken full effect? Therefore, for the purpose of this study, primary and secondary data were used. The secondary data was collected from existing literature, in specific journals, academic articles and companies' publications, in order to get knowledge about Insurance, Artificial Intelligence and the usage of this technology applied to the industry. The quantitative data was gathered based on industry indicators in Turkey, such as market shares and revenues, as well as information related to the main players in the Turkey market. From the Artificial Intelligence point of view, information regarding the investments made by enterprises in these innovations and the adoption of companies was collected through websites, studies and surveys previously done by other entities. In addition, a survey was held with Turkey residents, to assess respondents' opinion regarding the insurance industry, the personalization of Insurance policies based on AI methods and their perception of it [41].

3.1 ONLINE SURVEY

The online survey was an instrument to collect qualitative data to better understand Turkey respondents' perception about Insurance, their motivation for acquiring innovative insurance products and their opinion regarding the usage of AI for a personalized policy. The survey was expected to have a total of 100 respondents, residents in Turkey countries, with ages from 18 to 65+ years old and different occupations. It was shared through Social Media channels. The survey

was analyzed with excel, using descriptive statistics. The quantitative data from the study was collected through online searches on studies done for collecting information from the Insurance industry and the use of Artificial Intelligence. The information was extracted from the Insurance Turkey website, the Narrative Sciences survey in 2018 and the International Data Corporation.

3.2 SHORT TO MEDIUM TERM IMPACT OF ARTIFICIAL INTELLIGENCE IN INSURANCE ON TURKEY GOVERNMENT

Worldwide insurance premiums have been increasing but the state of the Turkey market registered a slow growth and weak performance during the last 5 years, with decreases in the annual gross written premiums between 2012 and 2017 of -0.5% for life insurance and -1,2% for non-life insurance. In 2018, the Turkey market corresponded to 31.6% of the total global insurance premiums, total premiums increased 6,2% and total claims and benefits paid grew 3,1% in comparison to 2017.

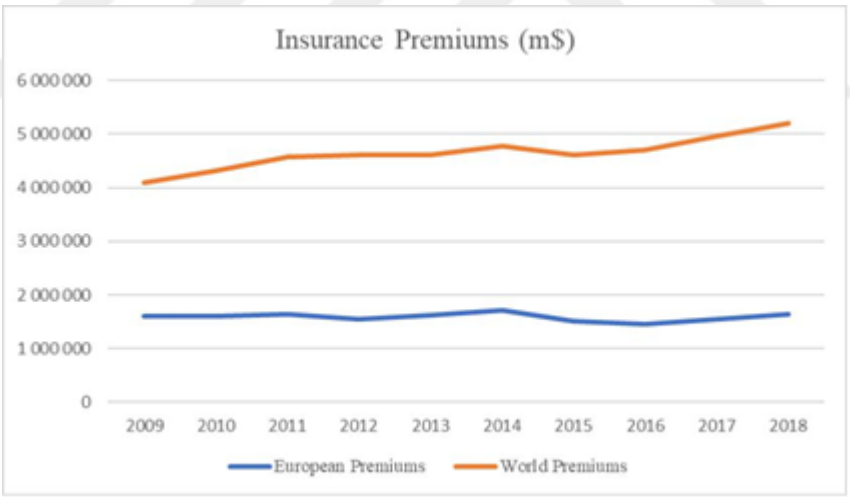


Figure 3.1: Source: Insurance Turkey 2019

Life Insurance was the segment with the highest premiums and claims paid. In Turkey, on average €2.170 was spent per capita in insurance, from which 1264€ was in Life insurance, 238€ in Motor, 232€ in Health and 174€ in Property.

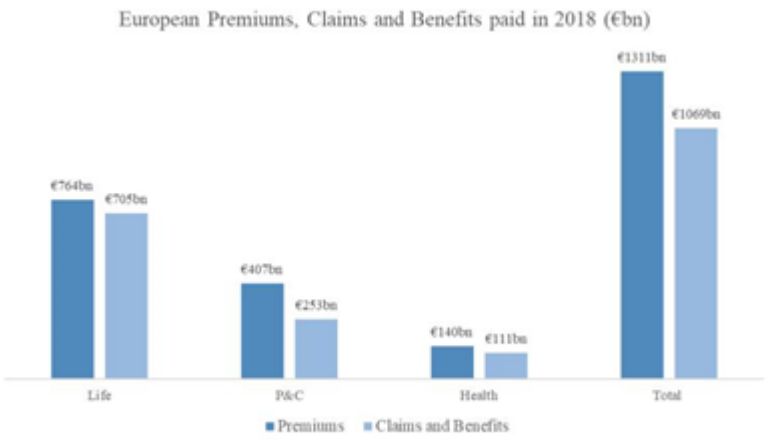


Figure 3.2: Turkey Premiums, Claims and Benefits paid in 2018

The countries with highest total premiums written in 2018 were the United Kingdom, France, Germany and Italy. Economic conditions had a strong effect on the performance of non-life insurances over the last years, and insurers must adapt to customer's expectations and innovate in operations by making use of Big Data, analytics and AI.

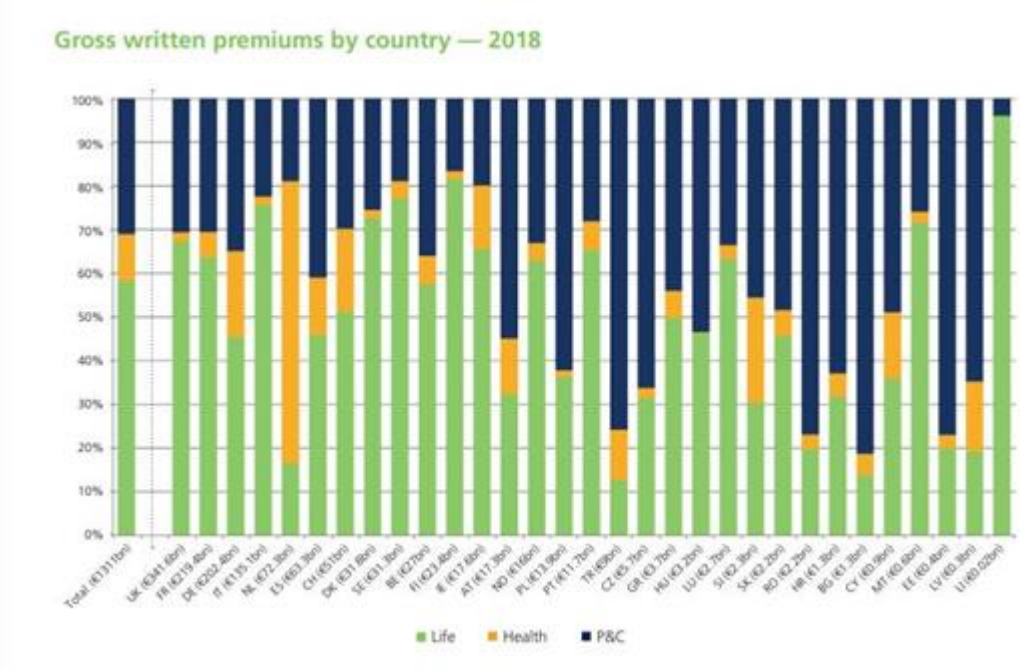


Figure 3.3: Source: Insurance Turkey 2019

3.3 ARTIFICIAL NEURAL NETWORK APPROACH

Artificial Neural Network algorithm is also a classification algorithm more commonly known as ANN. The ANN algorithm classifies the data into different groups which maximizes the utility for any researcher. Based on the different groups, we can separate the data belonging to one class from data belonging to another class. ANN has been applied to places such as time series prediction, face recognition to biological data processing. The ANN algorithm draws a hyper plane which is said to separate the two classes.

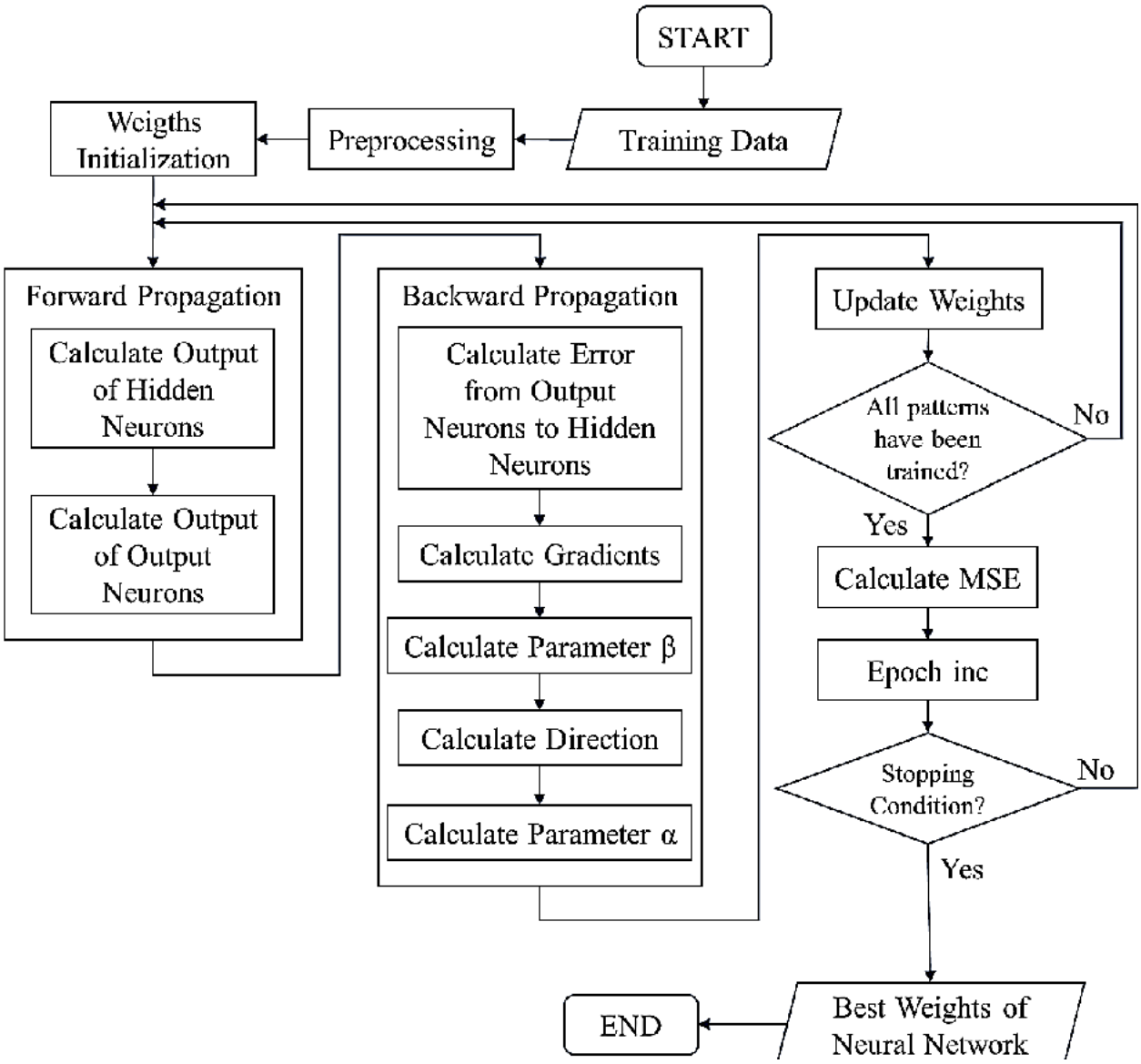


Figure 3.4: The flowchart represent the whole training process from input, updating weights, epochs and output.

The hyper plane is just not any ordinary line, but a line which tends to separate the two classes in such a way, such that the distance from the line to the closest point of both the classes remain as large as possible. In other words, it maximizes the distance from the hyper plane to each of the nearest data point of both the classes. The following equations describe the process on how this hyper plane is drawn and two different classes are identified.

$$w * x - b = 0 \quad (1)$$

The equation is of the hyperplane which separates the two classes equally. Here w is the normal vector to the hyperplane and x can vary for each set of data, b is the distance between the nearest point to the hyperplane and the hyperplane itself.

$$w * x - b >= 1 \, (2)$$

If, however the result to this equation is positive and equal to one, then we get the hyperplane of the positive class, and any value greater than 1 would give data points of the positive class.

$$w * x - b <= -1 \qquad (3)$$

Similarly, if the result is equal to -1, then it gives the hyperplane of the negative class and any value less than -1 would be for data points belonging to the negative class.

3.4 TRAINING WITH ARTIFICIAL NEURAL NETWORK

For our worldwide Governance indicator dataset, we first found the value of ‘k’ that best suited our case. We initially classified with ANN. However, this had no real logic behind it. As a result, we decided to find the optimal value that will give better and correct accuracy. It works by training our model on a range parameters and then finds out the best result based on comparison. In our case, we provided a range of values of k neighbors from 1 to 25 and checked which value of k neighbor provided the most accurate and reliable output. It is seen that best suited our data set. After determining the value, we load the data set. Duplicate rows with same values for all the four features are dropped to avoid training the same data over and over again. It also prevents testing and training with same dataset, avoiding accuracy levels that are unrealistic. Next, we split the data into train and test. 80 percent of the sample is train set and 20 percent of the data is test set. The 80 percent data is trained with ANN classifier with k initialized as 2. We then test the 20 percent

data. It is done by calculating the Euclidean distance between instances of the clusters or classes formed. The distances are then sorted in ascending order. The top N rows are then selected and the most frequent class of these rows are then returned. Finally, we calculated the accuracy and it was very high.

3.5 TRAINING, VALIDATION, AND TESTING SETS

All of the data is divided for the different purposes in this segment in which the portion of the data for the training will draw direct conclusions from the validation data. But If the accuracy of the validation set drops below the results of the training data, the network is starting to memorize the data it knows rather than generalizing on the concept.

This subset that will go through the process of testing is called testing data. In contrast to the validation set, it's only used once in the very end, to give a final score. The idea is that by building a model based on the validation results a certain amount of information bleed occurs, where the network will implicitly learn from the validation data. In order to avoid bad results, testing data is normally used as a reference for the performance.

- Training Set: 72% of the data
- Validation Set: 15% of the data
- Test Set: 13% of the data

3.5.1 Feature Scaling

When training neural networks, it is considered to be the best practice to scale the data. Otherwise different problems can be encountered such as degraded performance, unpredictable behavior of optimization algorithm and exploding gradient. There are different ways to scale the data, some of them rely on knowing the maximum and minimum values for the distribution, and as the data is not in some predefined range like, e.g. pixel value data, the standardization method is chosen, that relies on computing the mean and variance of the train set.

3.5.2 Feature Selection

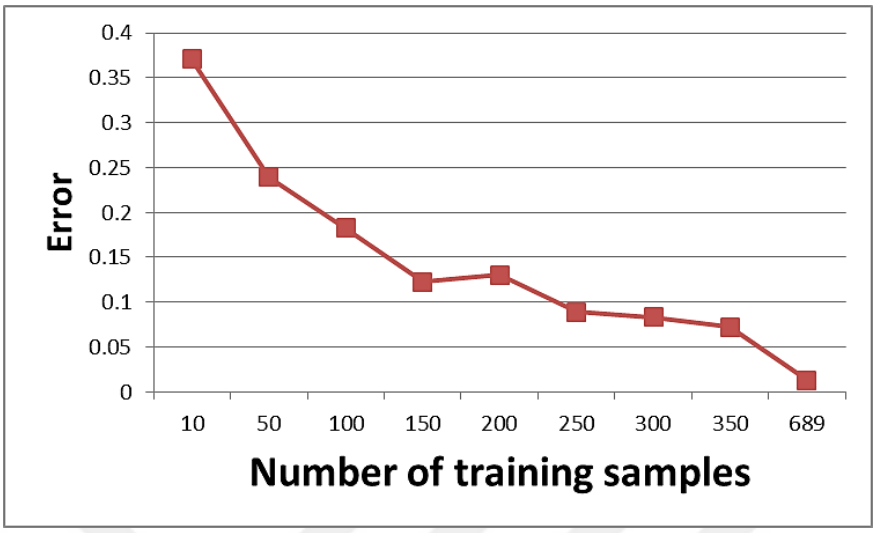
Some features of the data may be more important than others for the tasks of improving operational efficiency of government.By selecting a smaller number of features we will enable this work to be compared to other works and additionally demonstrate how neural network performance changes when applied to varying number of inputs.

4. RECOGNITION RESULTS

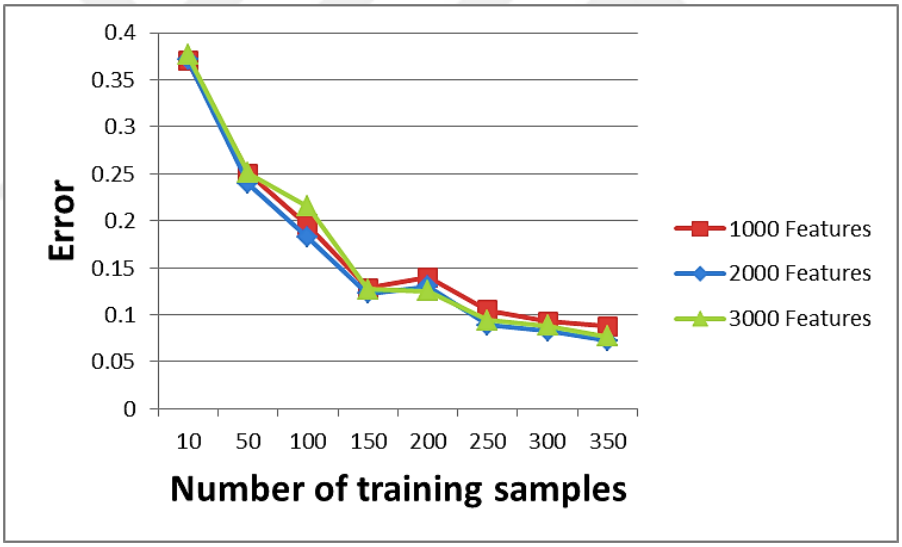
The results are presented in this section. The finest algorithm for operational efficiency of government and classification on the worldwide Governance indicator dataset is ANN. First of all, its accuracy is very high compared to other algorithms. Moreover, the execution time is also very low. As the complexity of the kernel decreases, the algorithm begins to get faster. The other algorithms are more or less same in terms of their execution time. Random Forest has the highest execution time due to the fact that it builds numerous tree before coming to a conclusion or decision. In this section, we have considered two parameters while analyzing the machine learning algorithm (ANN) on the dataset. They are accuracy in terms of percentage and time in terms of seconds. The experiment is carried using Matlab R2019a Windows environment on 32GB RAM and 2.8 GHz Intel Core i9. After carrying out several experiments using the artificial neural network, we have come to present all our findings here. We carried out artificial neural network algorithms on our data-set. We have done the test procedure with 20 percent data which is equal to 13583 data. All tests were run on a dual 12 core AMD Magny Cours equipped with a NVIDIA Tesla C2050 GPU with 14 Streaming Processors (32 cores each). The artificial neural network was created and trained on all 115 features. Default Keras optimization hyper-parameters were used. Training was limited to 100 epochs with additional condition of early stopping, using functionality provided by Keras. This ensures that the model is trained only until the score on validation set is not getting worse – this in turn helps to avoid overfitting the training set.

4.1 NUMBER OF TRAINING SAMPLES

The effect of varying the number of training samples. There is a clear trend of smaller test error as increase in the number of training samples. At 689 training samples, the error is only 1.32%. Increasing the number of training samples had the biggest impact in reducing the error when compared to other factors.



(a) Varying number of training samples.



(b) Varying the number of features. Number of features: 3000, and number of trees 3.

Figure 4.1: Experimental results being displayed in graphical representation.

4.2 NUMBER OF FEATURES

The effect of varying the number features. Across most of our experiments I observed that using 2000 features outperforms 3000 features, which outperforms 1000 features. However, the difference in error is small. This suggests that if I sampled 1000, 2000, and 3000 features again, I may observe a different trend. Overall, none of the alternative architectures achieved a higher

maximum classification accuracy on the five selected classes than the base multi-class design for budget balancing.

4.3 SYSTEM MODELING

The model trained in total for 24 epochs, taking about 27 seconds for each epoch, totaling in 648 seconds of training time. Figure 4.2 demonstrates the training curve. Loss is defined as MSE between original and predicted values. Training set had 15310 samples and optimization (validation) set 15311 samples.

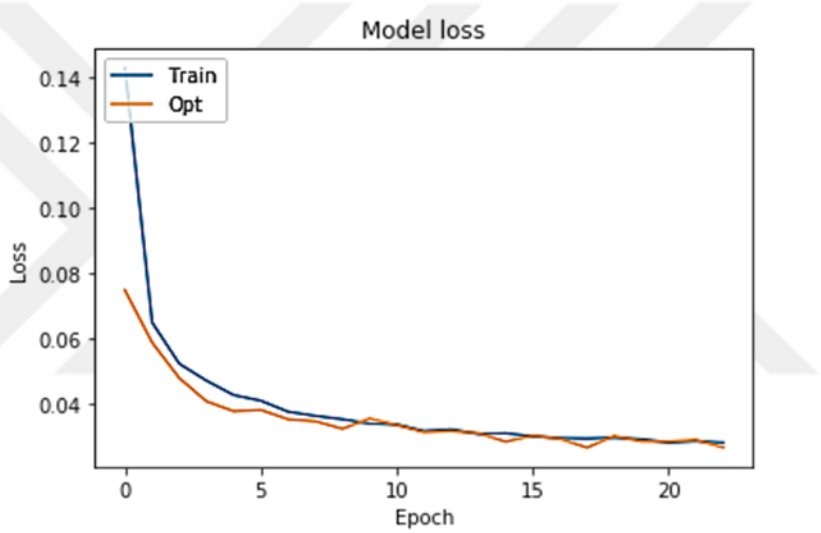


Figure 4.2: Artificial neural network training curve and optimization curve.

Additionally, neural network models were trained on reduced number of features for the purposes of their comparison. Some modifications in the training procedure were introduced, such as the maximum number of epochs was limited to 50.

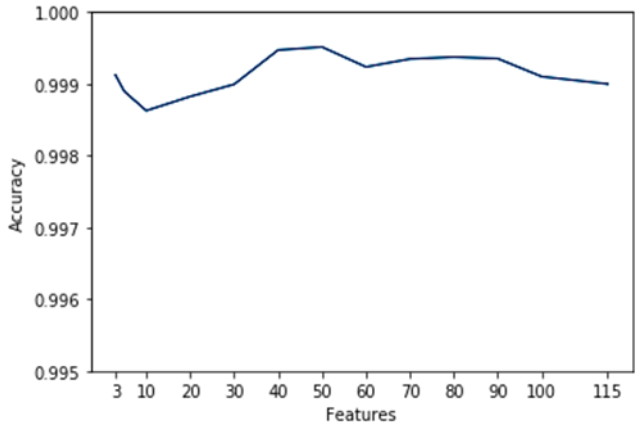


Figure 4.3: Neural network accuracy for different number of features.

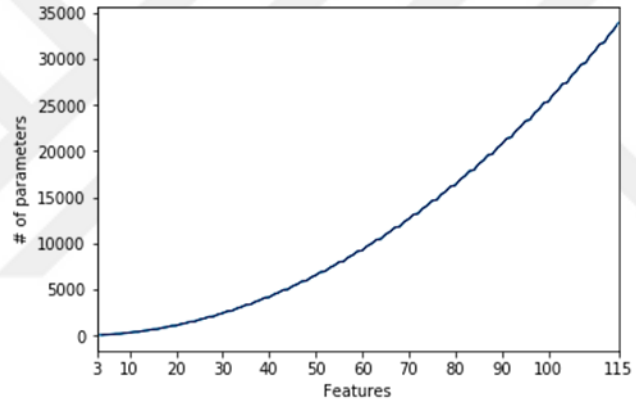


Figure 4.4: Artificial neural network number of parameters for different number of features.

We make a comparison between more traditional ML methods for operational efficiency and the artificial neural network implemented in this work.

4.4 PERSONALIZATION OF AN INSURANCE POLICY

Different suggestions associated with the personalization of an insurance policy were presented in the survey (Figure 4.5), and the option with highest interest for Turkey respondents was “I can activate/deactivate it whenever I need/want”, considered as Very Interesting, followed by the proposals of “Personalize the extent of services coverage” and “Define the amount I need for my coverage”. The option considered less interesting for customers was “Insurers using Artificial Intelligence to know me better”, considered Moderately Interesting (2.95).

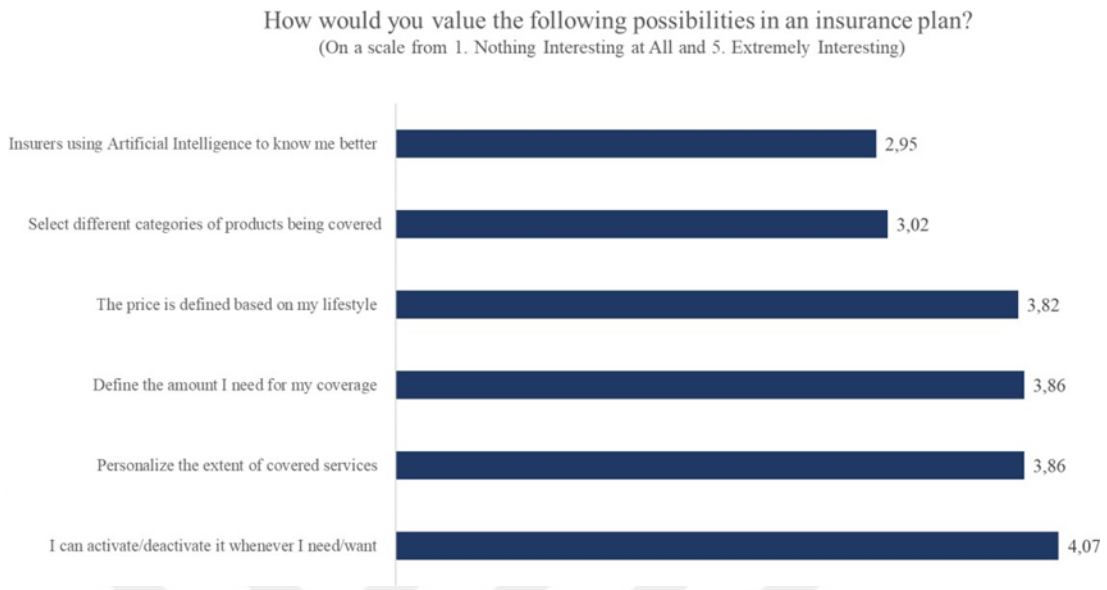


Figure 4.5: Suggestions for personalization of insurance

However, when respondents were asked about their level of agreement with the usage of Artificial Intelligence techniques to create personalized insurance policies, the majority of respondents agreed (35%), followed by a neutral opinion (28%).When asked to consider that a personalized policy made by Artificial Intelligence techniques would be able to define the services coverage based on their lifestyle, having what they need and at a price based on their purchasing power, but it required information about themselves (e.g. information shared online), the majority responded that they still agree (33%) or are neutral (26%) about it. However, negative responses for disagree (20%) and strongly disagree (14%) were also relevant, and it is relevant to highlight that the neutral opinions might be more likely to become a negative position towards this innovation[42].

4.4.1 Information Allowed to be used for Personalization of an Insurance Policy

When respondents were asked about if they would allow or not allow insurers to use certain information about themselves, the most ones that respondents would allow to be used is their physical activity, driving behavior, eating patterns and shopping behavior. The information that they would definitely not allow insurers to use would be photos of themselves and social media activity (Table 4.1). This information could be gathered mainly though their smartphones, devices for health monitoring, car sensors and fitness trackers (Table 4.1).

Table 4.1: Customer's information to be used in insurance

Information that Turkey respondents would allow to be used for personalization of their insurance (% of respondents)	
Would Allow	**Would Not Allow**
Physical activity (67%) Driving behavior (65%) Eating patterns (54%) Shopping behavior (41%)	Photos of themselves (87%) Social media activity (75%) Social life habits (67%) Other behaviors (e.g. Smoking) (51%)

Table 4.2: Devices for data collection

From the following devices, which ones would you allow to share information for AI processes? (% of respondents)	
Smartphone	55%
Devices for health monitoring (e.g. blood pressure wristbands)	52%
Car sensors	35%
Fitness trackers	34%
Computer	29%
Payment methods	24%
Smartwatch	22%
Clothing with sensors	14%

4.5 EVALUATION USING ARTIFICIAL NEURAL NETWORK

The confusion matrix for ANN. Out of the total data, 80 percent is used to train the data. The remaining data is used to test the data. From the figure, we can see that 8703 data are True positive. These are the data rows whose labels are labelled as normal or 1 and are also identified as normal or 1. Similarly 3868 data are true negative. These are data that are identified as abnormal or 0 and are actually labelled as 0 or abnormal. The 861 data which are actually abnormal or 0 are predicted to be normal or 1. Similarly, 421 data that are actually normal or 1 are predicted to be abnormal or 0. This gives ANN a higher accuracy rate compared to the other algorithms, which is about 95.2 percent. The reason for this high accuracy is due to the fact that the hyperplane that is drawn can separate the two classes accurately. We have used a sigmoid hyperplane to separate these data into their according classes for which the sigmoid line performs very well at separating the abnormal

and normal data. Other kernels of Linear lines will perform well because our data is very continuous rather than discrete. Since it is spread out everywhere randomly, we preferred to choose a sigmoid curve.

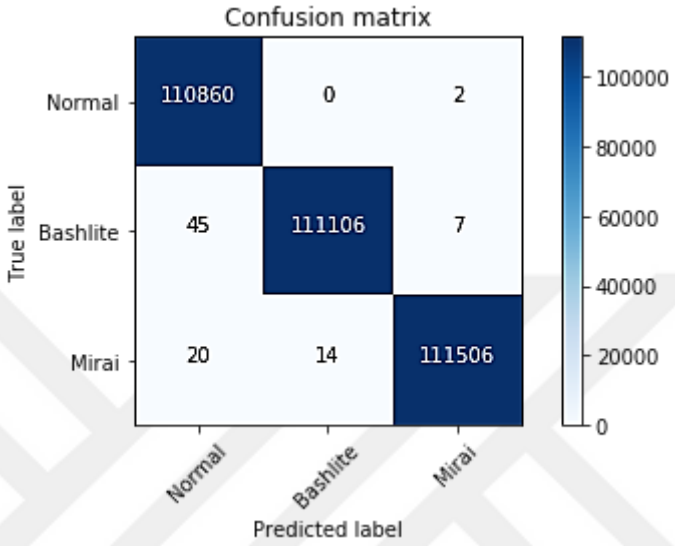


Figure 4.6: The confusion matrix for ANN.

Artificial neural network works better with multiple features as input. Since our data set has 4 classes of features, the algorithm works well. The accuracy also depends on how many classes the artificial neural network algorithm has to convert it to. In our case, the logistic function has to map between two classes. In binary artificial neural network implementation, the value set as the threshold value for the boundary between two classes can be high which means that it can separate between different sets of data more accurately. On the other hand, there would have been multiple thresholds for each set of output classes if there were more than 1 group of outputs. Thus, each threshold would have been lesser in value and the difference between one threshold value for one class and another threshold value of another class would also be lower. This can eventually lead into more errors. Thus, for binary artificial neural network, the chances of making mistakes is lower. Therefore, artificial neural network in our case has a high value of accuracy.

4.6 REGULATION APPLIED TO INSURANCE IN TURKEY

- **Regulation against discrimination**

The United Nations declaration of human rights banned discrimination based on identity (race, gender, nationality and others) and after 2014, it was stablished the directive 2004/113/EC for equal treatment of men and women, stating that the use of gender as an actuarial factor cannot influence individuals' premiums and benefits in insurance.

Although Turkey citizens have the right to equal treatment, the Turkey legal framework defined exceptions to this right, stating that financial products' prices (e.g. insurance) can vary by person based on age and health status, due to the Ex-ante profiling, that is related to actuarial factors . Premiums must be set when clients acquire the insurance and the health status can only be discriminated in the moment of acquisition, and it cannot be changed due to health aggravation or diseases acquired after it [43].

The Ex-ante profiling is a process that evaluates personal information to define a profile that corresponds to a specific level of risk based on probabilities, instead of on the actual costs incurred, and that the profile justifies the difference in the premium price (e.g. younger drivers paying more). The Ex-ante differences in insurance premiums are discriminative when they are based on identity characteristics and if differences in premiums allow some customers to receive more economic benefits than others. The Ex-post differences in the insurance premiums based on behavioural factors are discriminative when they are not calculated based on the actual damages incurred. The analysis of customers behaviours should be severely regulated in order to avoid discriminative treatments and exclusions and should be done to guarantee the minimum access to the insurance needed to participate equally in society (Finance Watch, 2019).Connected devices or cars used to innovate in insurance should be strictly used to allow discounts in premium prices and to guarantee the minimum insurance needed (Finance Watch, 2019)

- **Regulation for Data Privacy**

Insurers understand the importance of data protection, as data processing is an important part of their business, in order to analyse the risks that individuals want to cover and tailor their products

based on it, for evaluating and paying customer's claims and benefits and for detecting and preventing fraud.

In May 2018, the General Data Protection Regulation (GDPR) has been implemented in Turkey and insurers must consider it when making strategic decisions. The GDPR's concepts and requirements had to be clarified through guidelines and some of these have not been finalised yet, including the ones in regards to codes of conduct, that would facilitate insurer's compliance with GDPR with guidance on processing data and guarantees to consumers on how their data is being processed[44].

When it comes to the dualism between GDPR and innovation in insurance, the guidelines may prevent the development of innovative products based exclusively in automated techniques (e.g. real-time insurance through mobile phone apps).

Companies in Turkey have been limited by the GDPR for innovation, and if they want to create a personalized policy for their customers based on their data, according to this regulation, the data cannot have any type of PII (personal identifiable information) and documents cannot have anything that identifies customers as an individual, know what they are doing, in order to tailor their insurance based on it.

4.6.1 Research Question 1: What is the impact that AI will have in the Insurance industry in Turkey?

Regulation restricts innovation due to the burden and risk aversion that it has and there is a risk of regulators banning the use of technologies that they do not understand, leading insurers to limit innovations in areas prone to be legislated in the future. Yet, regulators in Turkey are trying to adapt to technological changes.

One of the most challenging regulation that has been applied in a for all industries was the General Data Protection Regulation, but despite this regulation, customers are willing to share data and insurers must focus on creating new flows of data and create data lakes to be used for improving their business . The GDPR law will not be limitative if consumers know what their data will be used for and give consent. Turkey respondents would be willing to share their data from their smartphones, devices for health monitoring, car sensors and fitness trackers, and a solution that

insurers could use to gather data with compliance with the GDPR could be creating new services and apps related to sports, health monitoring or driving behaviour, for instance, to complement their insurance and add value for customers.

Another limitation is regulation against discrimination in insurance. The EU law does not consider a price discriminative if the premiums are proportionate to the specific level of risk and the statistical data on which actuarial calculations are based is reliable (Finance Watch, 2019). However, according to the law, connected devices can only be used to guarantee discounts in premiums, and it can be a barrier for insurers to use it to personalize the whole policy.

Furthermore, the financial services sector in Turkey, including insurance, has been gradually deregulated by the Turkey to create a single Turkey market. In the future, this flexibility in regulation in EU countries could be applied in favour of the disruption of a market of one in insurance.

According to the survey results, Turkey customers would be willing to have personalized policies created by AI, and the top three categories that respondents would prefer for these innovative policies would be Health, House and Life . Turkey governments are receding from providing pensions, health insurance and other services, and the ratio of government spending to GDP decreased from 50.1% in 2009 to 45.8% in 2017. Furthermore, in the future, life insurance may have the opportunities from the retreating role of the state and the growing need for individuals to guarantee their own retirement savings and income. As a result, these changes can be an opportunity that Turkey insurers can tackle to innovate in their life and health insurance, with new offers for attracting new clients or improving its products for existing ones. Turkey insurers have been sluggish in adapting technology-enabled transformation and a solution may be partnering with InsurTechs for distribution and innovation. The impact of the disruption of policies created by AI would be an added value for Turkey policyholders and an opportunity for insurers to be more innovative and aligned with customers' expectations. However, Turkey insurers must bear in mind the influence of the existing regulations, especially the most recent GDPR, and the possibility of new ones, which may be the biggest threat for the full effect of this disruption in the Turkey market.

4.7 THE FUTURE OF THE INSURANCE INDUSTRY WITH THE IMPACT OF AI

4.7.1 The current state of Artificial Intelligence

Artificial Intelligence is the capacity of a machine to perform tasks similarly or better than a human. Worldwide investment made in AI is expected to reach $79.2 billion in 2022 with a 5 years CAGR of 38% between 2018-2022 and Western Turkey will be the region with highest investment in this technology. Industries worldwide are being disrupted by technological advances, which comprises solutions based on Artificial Intelligence, that are being implemented by 61% of the companies interviewed by the Narrative Science in their businesses to leverage new opportunities that digitalization brings, such as new business opportunities (Narrative Science, 2019)

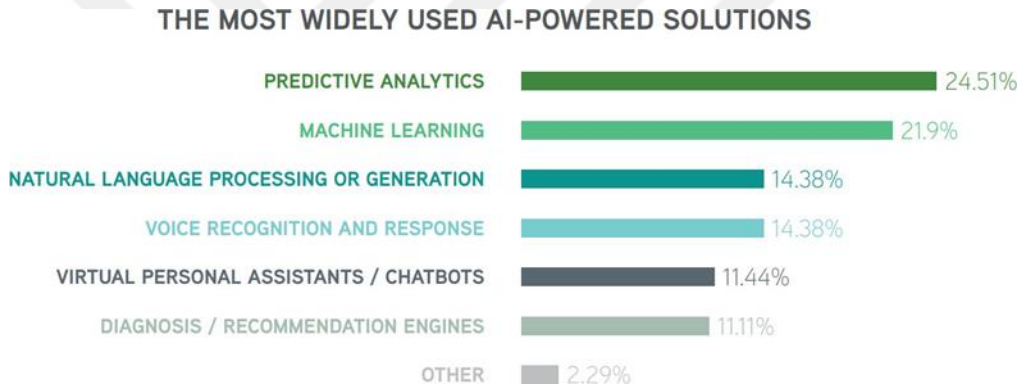


Figure 4.7: Source: Narrative Science 2019

The respondent companies exposed that the benefits from using AI in their businesses are for "identifying business opportunities", "automating repetitive tasks", "improving workforce productivity" and "competing with peers".

4.7.2 Insurance state and main challenges

The insurance industry is growing worldwide and the improves on the financial results from insurers have been supported by a sustained economic growth, increasing interest rates and higher investment income, which will be the tendency on the following years. The concept of insurance has been evolving over the years and insurers have an urge to rethink it, in order to keep aligned with the changes in customer's lifestyles and behaviours. This is conveying new risks and insights for insurers, incentivizing them to adapt their offers and rethink their business models. One of the

challenge in insurance is the longevity of population, that is increasing the number of diseases and the rise of costs in the health sector, which brings more expenses for insurers. Customers are also more concerned with being healthier and make more use of their health insurance. In addition, they are more conscious about the environment and living a sustainable life, changing their mobility behaviours with new solutions that are less harm (e.g. shared scooters) and with other means of transport to avoid using private cars (e.g. uber)

There are different motivations for acquiring and using an insurance. In the case of the job accidents and auto insurance, it is mandatory by law and policyholders hope not to use it, but in the case of health insurance, consumers expect to make use of it more often.

For younger customers, it is necessary to find new business models, as these generations want on demand services and show a tendency for being more conscious regarding where they spend their money. Hence, insurance must be accessible, make sense for their needs and add value. In insurance there are few points of contact and insurers observe a detachment from clients, resulting in a need to improve their service and increase engagement with customers (. Lastly, there has been a rise of new competitors from other sectors, other than financial ones, and it will bring one of the main changes in insurance, due to their disruptive business models. These competitors are native from the internet and digital services, have data from customers that can be leveraged and are willing to innovate in new services with shorter margins, to assess if they make sense for their business. All these new trends are resulting in new information for insurers, that will be a challenge when calculating premiums, as these are typically based on historical data.

4.7.3 Stakeholders in insurance

In Insurance, there are several stakeholders with different importance and involvement. Starting with the Primary Stakeholders, the most prevailing are the Insurance Companies, as their acts influence all the other stakeholders involved. Secondly, the end consumers (policyholders), that may change the way they consume insurance policies due to the changes in the industry. Thirdly, the Regulators, that regulate insurers in a national, and worldwide level. Fourthly, brokers and agents, as the way they sell insurance policies may change or be substituted. Lastly, the InsurTechs that came into the market with disruptive business models.

The secondary stakeholders are the employees, reinsurers, banks, firms and health providers.

4.7.4 How to create an insurance policy personalized by Artificial Intelligence - Technical Models

The creation of a personalized policy would be possible either by a Supervised Learning or Unsupervised Learning. The first one requires a training set based on past data with customer's characteristics and insures do not own it. The Unsupervised Learning would require going beyond the basic information that insurers own and having a rich characterization of the person. The hardest part is to understand what the person effectively likes and doing the match between the policy created and the person. There is no training set with this type of personalized policies to be used in the system, as insurers only sell a set of generic policies , and to tailor an offer based on a person's lifestyle and behaviours, it is necessary to have data and insurers would need to work in new actuarial and risk models .To create a personalized insurance policy, it requires a mechanism to generate policies with several parameters to create an infinite number of policies, an algorithm of parametrization to accept the generated policy and the person due to risk evaluation and create the training set to train the system .Algorithms work through the data and try to replicate the past, and to implement AI solutions on companies it is necessary to have data and some difficulties may be creating it and training the systems . As insurers do not own the training set, some suggestions to create it are:

1. **Doing questionnaires with customers and simulating the past:**

A way to create a training set would be simulating the past and doing questionnaires with customers to assess their preference from a set of possible policies. This way, it would be possible to create a personalized set of information to train the systems and for algorithms to learn from it, and insurers would know which policy a person with determined characteristics would prefer. If a person would choose its policy from an infinite possibility of policies, it could be used in the future for the system that would learn to offer policies to people.

2. **Create the training set with the help of agents and brokers:**

In order to assess what their customers would want based on agents and brokers' experience and make a hypothetic generation of people with certain policies to evaluate it.

In the future, AI could collect additional data from customers and do non-trivial data mining to obtain further information about customer's preferences. The key information is the type of policy that the person wants and is the most difficult to obtain.

3. Doing questionnaires exclusively for new policies:

A third option could be implementing this strategy only for new customers who want to acquire a personalized insurance. This way, it would be possible to create their profiles with information provided by themselves in a questionnaire and adapt the policy over the times with AI and the data that they share from their devices and apps.

4.7.5 Limitations of Artificial Intelligence

Nowadays it is possible to understand the content of a text through AI, but that is not possible to do causality and make a reasoning of it. Companies can only be better in detecting patterns, not relating them. For Natural Language Processing, the results in English are better than in other languages because it is an easier language in the grammatical point of view and because there exists more noted data.

4.7.6 Barriers for personalization of policies

- **Mutualism in insurance**

Mutualisation of risks is a key principle in insurance and any acts that reduce or limit it should be restricted to fair and relevant purpose and the differences in prices should incentivise less risky attitudes and increase the contribution of policyholders that had higher damages. The evaluation of risk of a person must be done in a cautious way, as it can result in policies that are impossible to afford for some consumers, because they need it and will have to cover their costs.

- **Bias and ethical issues**

Machine learning has each time more utility but it is crucial to think about the ethical questions and the bias that might be registered in data. Algorithms learn from this data and in order to avoid having biases embedded, insurers must decide if they want to use the data as it is or apply techniques to change it. To change bias, companies must get the algorithm, clean or remove some

data and add new false to balance, so some variables are not statistically significant. Having one to one policies can harm some clients due to their data, as their insurance may be better or worse based on it.

- **Data privacy**

One of the new strategic technology trends that will impact enterprises in 2020 (Appendix 12) is Transparency and Traceability, as the evolution of technology is originating a trust crisis, caused by customers becoming more aware of how their data is being collected and used and how companies are gathering and storing this data. Another issue comes associated with legislation such as the GDPR, that is having impact worldwide in organizations (Gartner, 2019).

- **Regulation in Insurance**

Regulation will be the main barrier for innovation in insurance with personalized policies created by AI. Firstly, for creating this kind of policy, it would require data from customers that may be hard to access due to the GDPR. Secondly, according to the Ex-ante moral hazard, the differences in premiums cannot depend on identity factors and the Ex-post states that differences in premiums cannot give economic benefits to some customers based on their behaviours (Finance Watch, 2019).

4.7.7 Possible scenarios

Three possible scenarios to evaluate the impact of the disruption of policies created by AI in the insurance industry were created based on players' reactions and past actions, secondary research done and calibrated with expert interviews insights. Scenarios were attributed equal weights and calibrated to assess its likelihood to happen.

5. DISCUSSION

Most Turkey residents own at least one type of insurance (84%) and the survey results shown that the categories with more policyholders are Health, Auto, House/Property and Life insurance. The main reason for not owning any insurance is because it is considered as too expensive (Figure 5.1).

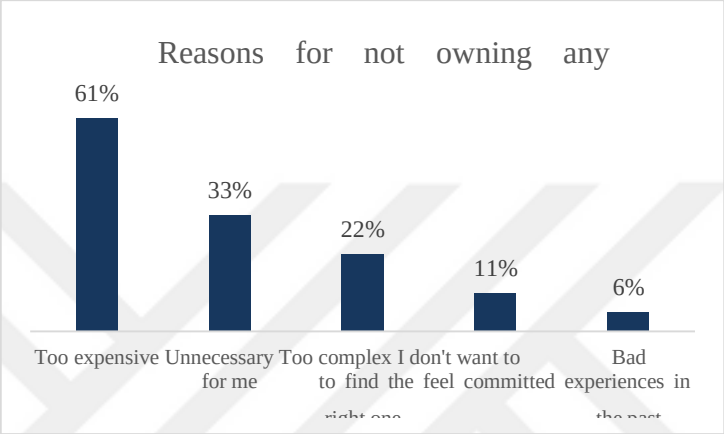


Figure 5.1: Reasons for not owning an insurance

The most important insurance categories for Turkey respondents were Health (4.58 out of 5), Auto (4.34) and House/Property (4.26). The categories Life (3.82), Personal Accident (3.62), Pets (2.93), and Technological Items (2.14) had lower scores. Furthermore, the insurance industry was considered by Turkey respondents as traditional, useful, apathetic, expensive, relevant, complex and more untrustworthy than trustworthy (Table 5.1).

Table 5.1: Characterization of the insurance industry

How would you describe the Insurance industry role in government efficiency?		
Classification	**Innovative**	**Traditional**
%	21%	79%
Classification	**Useful**	**Useless**
%	88%	12%
Classification	**Empathetic**	**Apathetic**
%	25%	75%
Classification	**Cheap**	**Expensive**

%	12%	88%
Classification	**Relevant**	**Irrelevant**
%	90%	10%
Classification	**Simple**	**Complex**
%	14%	86%
Classification	**Trustworthy**	**Untrustworthy**
%	49%	51%

Artificial neural network was also used and it works by using multi-variable analysis as it predicts a single outcome from multiple variable features The model is trained by using the neural function which sets a threshold value in probability that finally determines the two classes. This algorithm shows an accuracy of 95.2 percent in 2.17s. Artificial neural network is implemented on the network layer and input features that were trained for improving efficiency and allocating budget equally.

6. CONCLUSION AND FUTURE WORK

In this novel research work, we have developed an artificial intelligence-based system for improving efficiency of government, we have taken insurance into consideration. Insurance is a highly regulated industry and is considered as expensive, traditional and with low engagement by its customers. The response to the Research Question 1 is that the impact of AI policies in a Turkey level will be made more difficult. Turkey respondents would be willing to have this type of personalization and it could be a way to tackle the respondents who stated that they do not have an insurance because it is too expensive or unnecessary for them. However, due to the hurdle of the recent GDPR law, regulation will be the limitation that will be more difficult to mitigate. In addition, to create a policy using Artificial Intelligence, it is necessary to have a data set with information from customers that would be used to design the most appropriate policy, based on the customer's exact needs. For this, several strategies can be taken, such as doing questionnaires with customers or with the help of brokers. For Research Question 2, on a worldwide scale, the disruption of AI policies would be faster to be implemented. Traditional insurers will take longer to adapt this innovation due to the higher risk aversion related to regulation and the portfolio of existing customers, as the impact in revenues of innovations will be higher. However, partnering with Start-Ups and InsurTechs can be a strategy to improve and find new solutions to keep aligned with the demand for higher customization of polices. Additionally, it will be necessary to create protocols and data ecosystems to guarantee that data can be used and shared. Artificial Intelligence has still some limitations and for insurers working in English, it may be faster to implement a market of one strategy, due to the higher amount of noted data and because it is an easier language.

In the future, insurers must be able to rethink their mutualism models and strategies to stay competitive with the new entrants in the market. For this, it is essential to develop services and ecosystems that improve the engagement with customers and that will be the source to get access to customer's data. Telematics is already being used in car insurance, with a concept of "pay as you drive", in which premiums are based on the driving behaviour. It is a good indicator that other categories can move towards the same direction and use sensors to monitor customer's behaviours. InsurTechs also offer a concept of Usage-Based Insurance and micro coverage insurance for technological items that be activated/deactivated, for instance, being one of the possibilities that Turkey respondents value the most. A market of one in insurance policies can harm some clients

due to their data, and it can be an obstacle. However, in the future, insurers could develop to a more sophisticated model that would encompass the definition of the whole policy, but traditional insurers would need to move towards the same strategy and regulation would have to adapt to the changes that the advances in technologies and customer's expectations would bring. Moreover, ANN's run time and accuracy are very high in terms of improving the efficiency of the government. Hence, it can be included while improving operational efficiency. Furthermore, for the network layer, an artificial neural network has a very decent accuracy with the least execution time. Although we have achieved very good accuracy with the artificial neural network algorithms, there are other data sets with more features and different complexities where these algorithms might not detect intrusions with great precision. On those data sets, other machine learning algorithms or intrusion detection techniques might not work better. Lastly, the intrusion detection and classification is done on the end devices. We can carry out this detection and classification on the cloud or server in the near future with the help of deep learning as recommended.

In the future, the concept of insurance will change substantially and traditional insurers must adapt to the changes originated from the technological advances and in customer's needs and lifestyles. Insurance is a highly regulated industry and is considered as expensive, traditional and with low engagement by its customers. The response to the Research Question 1 is that the impact of AI policies will be made more difficult. Turkey respondents would be willing to have this type of personalization and it could be a way to tackle the respondents who stated that they do not have an insurance because it is too expensive or unnecessary for them. However, due to the hurdle of the recent GDPR law, regulation will be the limitation that will be more difficult to mitigate. In addition, to create a policy using Artificial Intelligence, it is necessary to have a data set with information from customers that would be used to design the most appropriate policy, based on the customer's exact needs. For this, several strategies can be taken, such as doing questionnaires with customers or with the help of brokers.

For Research Question 2, in a worldwide scale, the disruption of AI policies would be faster to be implemented. Traditional insurers will take longer to adapt this innovation due to the higher risk aversion related to regulation and to the portfolio of existing customers, as the impact in revenues of new innovations will be higher. However, partnering with Start-Ups and InsurTechs can be a strategy to improve and find new solutions to keep aligned with the demand for higher

customization of polices. Additionally, it will be necessary to create protocols and data ecosystems to guarantee that data can be used and shared. Artificial Intelligence has still some limitations and for insurers working in English it may be faster to implement a market of one strategy, due to the higher amount of noted data and because it is an easier language. In the future, insurers must be able to rethink their mutualism models and strategies in order to stay competitive with the new entrants in the market. For this, it is essential to develop services and ecosystems that improve the engagement with customers and that will be the source to get access to customer's data. Telematics is already being used in car insurance, with a concept of "pay as you drive", in which premiums are based on the driving behaviour. It is a good indicator that other categories can move towards the same direction and use sensors to monitor customer's behaviours. InsurTechs also offer a concept of Usage Based Insurance and micro coverage insurance for technological items that be activated/deactivated, for instance, being one of the possibilities that Turkey respondents value the most. A market of one in insurance policies can harm some clients due to their data, and it can be an obstacle. Public opinion is aware of these topics nowadays but if it is used for benefiting customers, it will be accepted. To conclude, the initial strategy should be doing adjustments in the premium paid based on customer's behaviours and with additional features for the insurance (e.g. possibility to deactivate it or change some coverages), because according to the current law, connected devices can only be used to benefit customers and reduce premiums. However, in the future, insurers could develop to a more sophisticated model that would encompass the definition of the whole policy, but traditional insurers would need to move towards the same strategy and regulation would have to adapt to the changes that the advances in technologies and customer's expectations would bring.

Limitations and Future Research

Some of the limitations regarding the execution of this study are related to the primary data generated. The survey conducted had a preponderance of Portuguese respondents in the sample and there is also a higher percentage of younger respondents. The survey might have also some positive bias, regarding the characteristics associated with the insurance industry and the ownership of insurance categories that may be different in some countries or age ranges. For future research, it could be done a survey more unbiased and with more representation of different countries. In addition, to assess the effect worldwide, it would also be interesting to do a survey

with different countries across the world. For future research, having experts with different nationalities would also be an added value. Moreover, doing a focus group with insurance customers and non-customers would be a way to present the idea of a personalized policy and understand the reaction and acceptance of customers in real time. Since it is an innovation in insurance, there is still not much content regarding using AI for personalizing an insurance policy, and there are more studies related to using this technology for claims processing, fraud detention or automatizing tasks.

REFERENCES

[1] J. K. Author, “Title of chapter in the book,” in *Title of His Published Book*, xth ed. City of Publisher, Country if not USA: Abbrev. of Publisher, year, ch. x, sec. x, pp. xxx–xxx.

[1] Davenport, T.H.; Harris, J.G. Competing on Analytics: The New Science of Winning; Harvard Business Press: Boston, MA, USA, 2007; pp. 3–22.

[2] LaValle, S.; Lesser, E.; Shockley, R.; Hopkins, M.S.; Kruschwitz, N. Big data, analytics and the path from insights to value. MIT Sloan Manag. Rev. 2011, 52, 21.

[3] Turkey Commission. Big Data Analytics for Policy Making; Turkey Commission: Brussels, Belgium, 2016; Available online: https://joinup.ec.europa.eu/sites/default/files/document/2016-07/dg_digit_study_big_data_analytics_for_policy_making.pdf .

[4] Jiang, H.; Shao, Q.; Liou, J.J.; Shao, T.; Shi, X. Improving the sustainability of open government data. Sustainability 2019, 11, 2388. [CrossRef]

[5] Flyvbjerg, B. Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again; Cambridge University Press: Cambridge, UK, 2011; pp. 55–60.

[6] Solesbury, W. Evidence Based Policy: Whence It Came and Where It’s Going; ESRC UK Centre for Evidence Based Policy and Practice, Queen Mary, University of London: London, UK, 2001; pp. 1–11.

[7] La Caze, A.; Colyvan, M. Evidence-Based Policy: Promises and Challenges. 2006. Available online: http: //colyvan.com/papers/ebp.pdf.

[8] Watts, R. Truth and politics: Thinking about evidence-based policy in the age of spin. Aust. J. Publ. Admin. 2014, 73, 34–46. [CrossRef]

[9] Mulgan, G. Government, knowledge and the business of policy making. In Proceedings of the National Institute of Governance Conference, Canberra, Australia, 23–24 April 2013.

[10] Duan, Y.; Edwards, J.S.; Dwivedi, Y.K. Artificial intelligence for decision making in the era of Big Data—Evolution, challenges and research agenda. Int. J. Inf. Manag. 2019, 48, 63–71.

[11] Sutcliff, S.; Court, J. Evidence-Based Policymaking: What Is It? How Does It Work? What Relevance for Developing Countries? Overseas Development Institute: London, UK, 2015.

[12] Jennings, E.T., Jr.; Hall, J.L. Evidence-based practice and the use of information in state agency decision making. J. Public Adm. Policy Res. 2011, 22, 245–266. [CrossRef]

[13] OECD. Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information; OECD: Paris, France, 2008; Available online: https://legalinstruments.oecd.org/public/doc/122/122.en.pdf

[14] Davis, P. Is evidence government possible? Paper presented at the 4th Annual Campbell Collaboration Colloquium, Washington, DC, USA, 19 February 2014.

[15] Triantafillou, P. The political implications of performance management and evidence-based policymaking. Am. Rev. Public Adm. 2015, 45, 167–181.

[16] Ferrandino, J. The enemy of teaching evidence-based policy: The Powell-Bush doctrine of public affairs. J. Public Aff. Educ. 2014, 20, 73–89.

[17] Jost, J.T.; Federico, C.M.; Napier, J.L. Political ideology: Its structure, functions, and elective affinities. Annu. Rev. Psychol. 2009, 60, 307–337.

[18] Esty, D.; Rushing, R. The promise of data-driven policymaking. Issues Sci. Technol. 2007, 23, 67–72.

[19] Bradley, C.J.; Penberthy, L.; Devers, K.J.; Holden, D.J. Health services research and data linkages: Issues, methods, and directions for the future. Health Serv. Res. 2010, 45, 1468–1488.

[20] Curtis, L.H.; Brown, J.; Platt, R. Four health data networks illustrate the potential for a shared national multipurpose big-data network. Health Aff. 2014, 33, 1178–1186. [CrossRef]

[21] Zhang, Y.; Xu, Y.; Shang, L.; Rao, K. An investigation into health informatics and related standards in China. Int. J. Med. Inform. 2007, 76, 614–620.

[22] Chen, H.; Wang, F.Y.; Zeng, D. Intelligence and security informatics for homeland security: Information, communication, and transportation. IEEE Trans. Intell. Transp. 2004, 5, 329–341.

[23] Chen, H.; Wang, F.Y. Guest editors' introduction: Artificial intelligence for homeland security. IEEE Intell. Syst. 2005, 20, 12–16.

[24] Ku, C.H.; Leroy, G. A decision support system: Automated crime report analysis and classification for e-government. Gov. Inf. Q. 2014, 31, 534–544.

[25] Sachs, J.D.; Schmidt-Traub, G.; Mazzucato, M.; Messner, D.; Nakicenovic, N.; Rockström, J. Six Transformations to Achieve the Sustainable Development Goals. Nat. Sustain. 2019, 2, 805–814.

[26] Oh, S.; Kim, S. R&D support for SMEs: Current status and performance analysis. STEPI Insight. 2018, 224, 1–30.

[27] Yang, H.B. Status of technology commercialization of SMEs and policy tasks. J.S.F. 2017, 37, 23–45.

[28] Presidential Advisory Council on Science and Technology (PACST). National Innovation Plan; PACST: Seoul, Turkey, 2018.

[29] Kohonen, T. Essentials of the self-organizing map. Neural Netw. 2013, 37, 52–65.

[30] Liu, Y.; Weisberg, R.H.; Mooers, C.N. Performance evaluation of the self-organizing map for feature extraction. J. Geophys. Res. 2006, 111, C05018.

[31] Reusch, D.B.; Alley, R.B.; Hewitson, B.C. North Atlantic climate variability from a self-organizing map perspective. J. Geophys. Res. 2007, 112, D02104.

[32] Mangiameli, P.; Chen, S.K.; West, D. A comparison of SOM neural network and hierarchical clustering methods. Eur. J. Oper. Res. 1996, 93, 402–417.

[33] Rutkowski, L.; Jaworski, M.; Pietruczuk, L.; Duda, P. The CART decision tree for mining data streams. Inform. Sci. 2014, 266, 1–15.

[34] World Customs Organization. WCO Customs Risk Management Compendium; WCO: Brussels, Belgium, 2011.

[35] Cooper, R.B.; Zmud, R.W. Information Technology Implementation Research: A Technological Diffusion Approach. Manag. Sci. 1990, 36, 123–139.

[36] Welch, E.W.; Pandey, S.K. E-government and Bureaucracy: Toward a Better Understanding of Intranet Implementation and its Effect on Red Tape. J. Public Adm. Res. Theory 2006, 17, 379–404.

[37] Welch, E.W.; Feeney, M.K. Technology in Government: How Organizational Culture Mediates Information and Communication Technology Outcomes. Gov. Inf. Q. 2014, 31, 506–512.

[38] Moon, M.; Bretschneider, S. Does the Perception of Red Tape Constrain IT Innovativeness in Organizations? Results from Simultaneous Equation Model and Implications. J. Public Adm. Res. Theory 2002, 11, 327–352.

[39] World Customs Organization. Customs and the UN 2030 Agenda for Sustainable Development. Available online: http://www.wcoomd.org/ru-ru/topics/key-issues/customs-in-support-of-sdgs.aspx

[40] World Customs Organization. Customs' Role in Achieving Sustainable Development Goals; WCO: Brussels, Belgium, 2018.

[41] Luna-Reyes, L.F.; Gil-Garcia, J.R. Digital Government Transformation and Internet Portals: The Co-evolution of Technology, Organizations, and Institutions. Gov. Inf. Q. 2014, 31, 545–555.

[42] Bozeman, B. Bureaucracy and Red Tape; Prentice Hall: Upper Saddle River, NJ, USA, 2000.

[43] Kim, H.J.; Pan, G.; Pan, S.L. Managing IT-enabled Transformation in the Public Sector: A Case Study on E-government in South Korea. Gov. Inf. Q. 2007, 24, 338–352.

[44] Gibreel, O.; Hong, A. A Holistic Analysis Approach to Social, Technical, and Socio-Technical Aspect of E-Government Development. Sustainability 2017, 9, 2181.