



**REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKES SCIENCE AND TECHNOLOGY
UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING**

**DEMAND FORECASTING USING ARTIFICIAL NEURAL NETWORKS
FOR POWER TRANSFORMERS**

**DOĞUKAN GÖRÜR
MASTER OF SCIENCE**



**REPUBLIC OF TURKEY
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY
UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING**

**DEMAND FORECASTING USING ARTIFICIAL NEURAL NETWORKS
FOR POWER TRANSFORMERS**

**DOĞUKAN GÖRÜR
MASTER OF SCIENCE**

**SUPERVISOR
ASSOC. PROF. DR. MUSTAFA GÖÇKEN**

ADANA 2022

I hereby declare that all information in this thesis 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 information that is not original to this work.

[Signature]

Doğukan GÖRÜR

ABSTRACT

DEMAND FORECASTING USING ARTIFICIAL NEURAL NETWORKS FOR POWER TRANSFORMERS

Doğukan GÖRÜR

Department of Industrial Engineering

Supervisor: Assoc. Prof. Dr. Mustafa GÖÇKEN

November 2022, 133 pages

Beholding the historical processes, the human instinct to predict the future has always been dominant. In the modern era, scientific and mathematical models have given civilizations an important tool in predicting the future. By the last quarter of the 20th century, artificial neural network studies gained momentum and gave a new perspective to forecasting techniques. Demand forecasting is one of the forecasting issues that come to the fore with industrialization. In this study, a demand forecast for power transformers has been carried out with the multilayer perceptron method, a type of artificial neural network. The factors affecting the demand have been determined by taking an expert opinion from a regional producer, and the data covering past 40 years regarding these factors were obtained from official sources. The artificial neural network created during implementation has produced consistent results at the rate of 97.99%. It was observed that it outperformed the ARIMA model created in the study for comparison. It was concluded that the multi-layer artificial neural network created in the study is a useful tool for demand forecasting of power transformers.

Keywords: artificial neural network, multi-layer perceptron, demand forecasting

ÖZET

YAPAY SİNİR AĞLARI İLE GÜÇ TRANSFORMATÖRLERİ İÇİN TALEP TAHMINİ

Doğukan GÖRÜR

Endüstri Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Mustafa GÖÇKEN

Kasım 2022, 133 sayfa

Tarihi süreçlere göz atıldığında insanoğlunun geleceği öngörebilme dürtüsü her zaman baskın gelmiştir. Modern çağlara gelindiğinde ise bilimsel ve matematiksel modeller geleceği öngörebilme bağlamında medeniyetlere önemli bir araç kazandırmıştır. 20. yüzyılın son çeyreğine gelindiğinde ise yapay sinir ağları çalışmaları hız kazanmış ve öngörüm tekniklerine yeni bir bakış açısı kazandırmıştır. Talep tahmini ise endüstriyelleşme ile birlikte ön plana çıkan öngörüm konularından biridir. Bu çalışmada ise yapay sinir ağlarının bir türü olan çok katmanlı algılayıcı yöntemiyle güç transformatörlerine yönelik bir talep tahmini gerçekleştirilmiştir. Talebi etkileyen faktörler bölgedeki bir üreticiden uzman görüşü alınarak belirlenmiş, bu faktörlere ait 40 yıllık veriler ise resmi kaynaklardan elde edilmiştir. Uygulama neticesinde oluşturulan yapay sinir ağı %97,99 oranında tutarlı sonuçlar vermiş ve çalışmada ayrıca oluşturulan ARIMA modeliyle karşılaştırıldığında daha üstün bir performans ortaya koyduğu gözlemlenmiştir. Çalışmada oluşturulan çok katmanlı yapay sinir ağının güç transformatörlerinin talep tahmininde kullanılabilir bir araç olduğu sonucuna varılmıştır.

Anahtar Kelimeler: yapay sinir ağları, çok katmanlı algılayıcı, talep tahmini



ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Assoc. Prof. Dr. Mustafa GÖÇKEN for his indulgent, benevolent, and gracious attitude throughout my thesis. Besides, I would also like to thank the thesis committee for sparing their valuable time and opinions.

I would like to thank my family, who played a big role in helping me get to where I am today and did not spare anything about my education and my future. Foremost, would like to thank my wife my candle of hope for her continuous support who always believed in me and encouraged me in tough times

TABLE OF CONTENTS

TABLE OF CONTENTS	viii
LIST OF FIGURES	x
LIST OF TABLES	xi
NOMENCLATURE	xii
1. INTRODUCTION	1
1.1. An Overview of Forecasting.....	3
1.2. Power Transformers and Transformer Theory	7
1.3. Transformer Market	11
1.4. Factors Affecting Demand of Power Transformers	15
1.4.1. Production Volume.....	15
1.4.2. Country Population.....	17
1.4.3. Total Power Installed	18
1.4.4. Total Power Generated	20
1.4.5. Consumption of Energy	22
1.4.6. Number of Enterprises	24
1.4.7. Electricity Consumption of The Topmost Exported Countries.....	25
1.4.8. Population of The Topmost Exported Countries	29
1.4.9. Number of Installed Power Transformers	33
1.5. Demand for Power Transformers.....	34
2. LITERATURE REVIEW	37
2.1. Demand Forecasting Applications from Various Topics	37
2.2. Forecasting Applications with Artificial Neural Networks	41
2.3. Applications of Artificial Neural Networks in Energy Sector	47
2.4. Hypothesis	52
3. METHODOLOGY	53

3.1. Historical background of Artificial Neural Networks	53
3.2. Concurrence with Neurology	56
3.3. Neural Network Architecture	57
3.4. Multilayer Perceptron.....	64
3.5. Training Processes and Learning Strategies	66
3.5.1. Training Processes of Multilayer Perceptron	68
3.6. Evaluation Principle of Outputs	77
3.7. Implementation	79
3.7.1. Multilayer Perceptron Application via IBM SPSS.....	79
3.7.2. Improving the Performance of the Neural Network	87
3.7.2.1. Modifying Number of Hidden Layers	87
3.7.2.2. Modifying Number of Hidden Neurons.....	87
3.7.2.3. Modifying Initial Learning Rate.....	89
3.7.2.4. Modifying Momentum Coefficient.....	90
4. RESULTS	93
4.1. A Comparison.....	102
5. CONCLUSION.....	104
6. RECOMMENDATIONS.....	106
REFERENCES	107

LIST OF FIGURES

Figure 1.1 Framework for forecasting and planning	4
Figure 1.2 Information flow in Sales Forecasting	5
Figure 1.3 Characteristics of Forecasting Methods and Their Relationship.....	6
Figure 1.4 Utilization of the power transformers in different phases	8
Figure 1.5 Transformer in no-load condition.....	10
Figure 1.6 List of importing markets for a product exported by Turkey part (Product:85)	13
Figure 3.1 Biological Neuron.....	56
Figure 3.2 Basic Neural Network Diagram	57
Figure 3.3 Sigmoid Function	58
Figure 3.4 Four-Layer Network Containing Three Bias Neurons	60
Figure 3.5 Sigmoid Curves without Bias Neuron	61
Figure 3.6 Sigmoid Curves with Bias Neuron	62
Figure 3.7 The conjunction AND.....	63
Figure 3.8 Multilayer Perceptron Network Diagram	64
Figure 3.9 Information Flow in an MLP Network	69
Figure 3.10 Local and Optimum Solutions Considering Weights	75
Figure 3.11 Flow Chart of Feedforward-back propagation Algorithm	76
Figure 3.12 Data View on SPSS	79
Figure 3.13 Configuration of Variables.....	80
Figure 3.14 Setting Partitions.....	81
Figure 3.15 Architectural Parameters.....	82
Figure 3.16 Training Parameters.....	83
Figure 3.17 Output Options	84
Figure 3.18 Saving and Exporting Options	85
Figure 3.19 Other Options	86
Figure 4.1 Final Network Diagram	94
Figure 4.2 Sales-Predicted Value	96
Figure 4.3 Predicted Value-Residual.....	96
Figure 4.4 Normalized Importance Chart	98
Figure 4.5 Comparison of Actual Demand with ANN and ARIMA Models	103

LIST OF TABLES

Table 1-1 List of products exported by Turkey Product:85.....	12
Table 1-2 List of exported products for Product 85	14
Table 1-3 Annual production volume of PT	15
Table 1-4 Address-Based Population of Turkey	17
Table 1-5 Total Power Installed by Power Plants	19
Table 1-6 Gross Generation of Energy	20
Table 1-7 Net Energy Consumption.....	22
Table 1-8 Number of Enterprises	24
Table 1-9 Electricity Consumption of The Topmost Exported Countries.....	26
Table 1-10 Total Electricity Consumption of Topmost Exported Countries.....	28
Table 1-11 Populations of The Topmost Exported Countries	30
Table 1-12 Total Population of The Topmost Exported Countries.....	31
Table 1-13 Number of Installed Power Transformers.....	33
Table 1-14 Historical Sales Amount of Power Transformers.....	35
Table 3-1 Commonly Used Activation Functions	59
Table 3-2 Performance Metrics.....	78
Table 3-3 Results with Default Parameters	86
Table 3-4 Results Improved with Number of Hidden Layers and Neurons	88
Table 3-5 Results Improved with Initial Learning Rate	90
Table 3-6 Results Improved with Momentum Coefficient.....	91
Table 3-7 Model Summary	92
Table 4-1 Parameter Estimates.....	95
Table 4-2 Independent Variable Importance Analysis	97
Table 4-3 Actual Demand vs Forecasted Demand	98
Table 4-4 Performance Metrics of Neural Network.....	100
Table 4-5 ARIMA Model Results	102

NOMENCLATURE

AC	: Alternating Current
AE	: Absolute Error
ANN	: Artificial Neural Network
ARIMA	: Autoregressive Integrated Moving Average
DC	: Direct Current
e.m.f	: Electromotive Force
GNP	: Gross National Product
LDT	: Liquid Dielectric Transformers
m.m.f	: Magnetomotive Force
MAD	: Mean Absolute Deviation
MAE	: Mean Absolute Error
MAPE	: Mean Absolute Percentage Error
MLP	: Multi-layer Perceptron
MLR	: Multiple Linear Regression
MNR	: Multiple Nonlinear Regression
MSE	: Mean Squared Error
PT	: Power Transformer
RMSE	: Root Mean Square Error
SME	: Small-to-Medium Enterprise
SSE	: Sum of Squared Error
SSR	: Sum of Squares Regression

1. INTRODUCTION

Failure is success if we learn from it

- Malcolm Forbes

Throughout history, humankind had always wondered about the future and its reflections have found a place in many fields from science to art. For instance, famous British futurist author H. G. Wells invented “Time Machine” and in the same vein American producer R. Zemeckis have made traveled many audiences in various dimensions of time with his “DeLorean”. Albert Einstein, on the other hand, went one step beyond the imagination with his relativity theory stating that the time travel is possible in another manner based on the fact that the time is connected with space, and the faster you go, the less you experience the time (NASA, 2020).

One of the biggest factors underlying this curiosity of human beings is undoubtedly the positive aspects of a successful foresight that can affect today's zero point. One of these positive aspects undeniably includes economic factors. In the fields mentioned above, this factor was also frequently observed, and outside of these areas, for example in science quantitative and qualitative methods were used to predict the future more successfully, taking the economic concerns into account. It would not be wrong to say that a successful future forecast can positively affect an individual at the zero point of time, as well as a legal entity. To be specific, an accurate forecast made with the outputs of a business has the potential of making this business more profitable or less likely to lose.

Prediction or forecasting techniques can be referred to as qualitative and quantitative forecasting techniques. It can be said that the prediction techniques, which could be grouped as qualitative and quantitative forecasting technique, to be used generally depends on the type of data to be used in practice. In practice, it is very difficult to adapt data to certain patterns. Artificial neural networks, as a quantitative technique, do the opposite by adapting itself to the model in practice thanks to its hidden layers and acquire a data-driven identity. Considering this data-driven approach, it is planned to use artificial neural networks as a forecast technique in this research. (Eğrioğlu et al., Yolcu & Baş, 2019; Lee, Song, & Mjelde, 2008)

Just as the forecasting method to be applied is important, the type of product to which the prediction will be applied plays an important role in the forecasts. The estimation to be made can also be related to whether the output is a product of a service system or manufacturing system. Quantitative demand forecasting methods are more applicable as output is generally concrete in manufacturing systems, however the situation is different in service systems as the product is often intangible in those systems and forecasting to be applied in any part of this system usually made with qualitative estimation techniques (Demirtaş & Anagün, 2006).

In this research, a demand forecast will be made for power transformers, which are associated with the electrical machine product family and the output of the manufacturing system, via a quantitative forecasting technique using an artificial neural network. As will be supported by the data in the later stages of the study, the most important reasons for selecting this product group as the study object are the fact that electrical systems immensely require this product in many areas such as generation, transmission and distribution, and the high trade, import and export amounts of it in both global and regional scale.

In the first part of the study, power transformers were defined, and their working principles are briefly mentioned along with their parts and utilization areas. Additionally, the market of power transformers has been presented in order to underline the significance of product by the trade amounts such as import and exports amounts which gathered in both worldwide and local scale. Thenceforth, the factors affecting the demand for power transformers were determined by obtaining an expert opinion and the data belonging to these variables were presented in detail. Finally, in the first chapter, the significance and requirement for forecasting have been stated together with the fundamental background information. On second chapter literature review have been performed together with the presentation of hypothesis. Subsequent chapter contains detailed information about the approach utilized in this study. In chapter five results have been proposed along with a comparison. And lastly on following two chapters while making a conclusion about the study, the situations encountered in this study are mentioned in order to shed light on the next studies together with discursive statements.

1.1. An Overview of Forecasting

Forecasting has been defined as “a statement about what will happen in the future based on the information that is available now” by Oxford Language (Hornby, 2000). In many areas from ancient times to modern times, it has found a place for itself in the lives of mankind (Hyndman & Athanasopoulos, 2013). In modern times, foresight is mostly used for commercial, economic, industrial and scientific purposes (Chatfield, 2000). Undoubtedly, one of the biggest reasons for its frequent use is the motive to minimize the uncertainty that exists between the upcoming events and the current time. (Makridakis et al., 1997).

Forecasting is used in several areas and is applicable to various time intervals. For instance, a good forecast maybe used to determine the need for a new power plant in the next five years or how many employees are needed for a week ahead. (Hyndman & Athanasopoulos, 2013). It is even used to predict a few seconds later in cases such as telecommunication routing or electrical utility loading (Makridakis et al., 1997). Intrinsically, forecasting is an important and effective tool in businesses regardless of the circumstances, usage area or time frame (Hyndman & Athanasopoulos, 2013).

Although forecasting is a statistical tool used in the short, medium and long-term time intervals from production to scheduling from transportation to personal affairs, it is a term that is often confused with planning (Hyndman & Athanasopoulos, 2013). While planning attempts to understand what the environment should be like, forecasting attempts to reveal how it will be. (Armstrong, 2002). The relationship between planning and forecasting could be illustrated as follows;

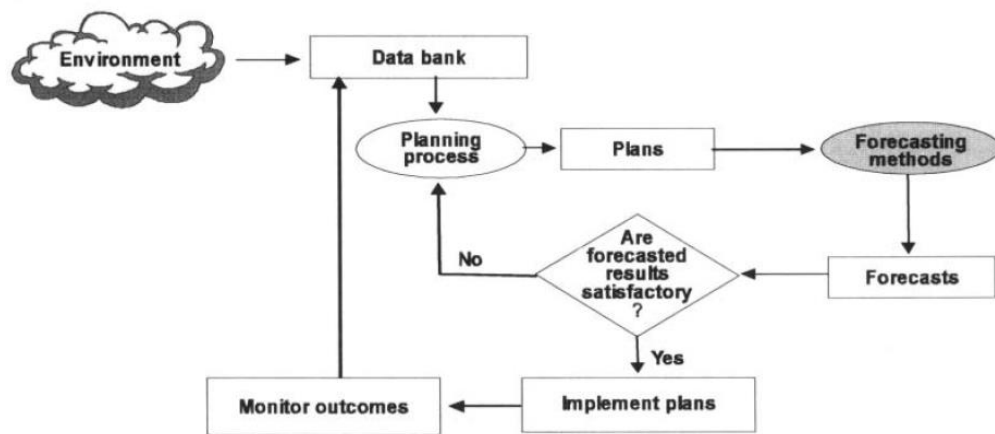


Figure 1.1 Framework for forecasting and planning (Armstrong, 2002)

Forecasting could be used in many departments where planning or foresight can yield positive gains such as sales, inventory or stock control, long term economic planning, production, capacity planning, budgeting, risk management, etc. (Chatfield, 2000). Forecasting should play a complementary role at the administrative level. Because in order to acquire a reliable forecast, there must be a sound information exchange and relationship between departments, otherwise some negative consequences may occur. For example, an unsuccessful sales forecast can lead to a chained misalignment of operating costs, budget, stock level, pricing, and cash flow (Hyndman & Athanasopoulos, 2013; Makridakis et al., 1997). The information flow mentioned above could be illustrated as follows since it underlines the impact of forecasting in overall.

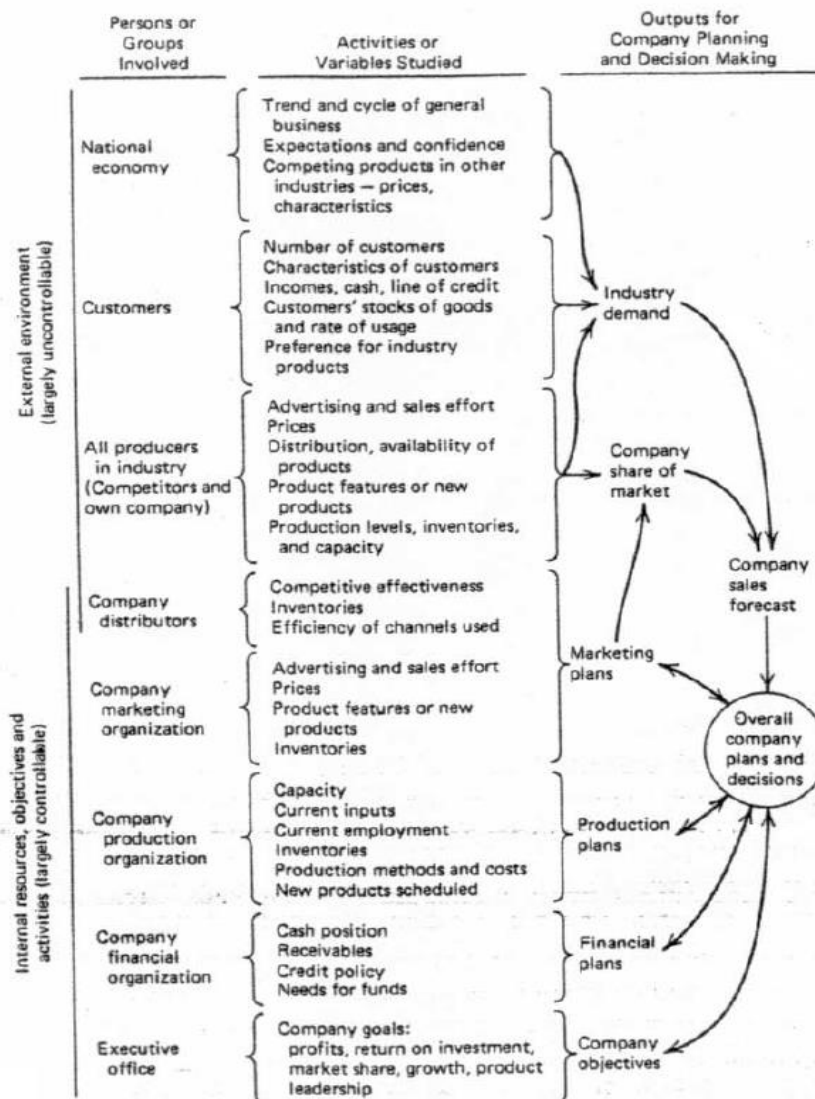


Figure 1.2 Information flow in Sales Forecasting (Makridakis et al., 1997)

It could be noted that choosing the appropriate forecasting method generally depends on the type and amount of data available (Hyndman & Athanasopoulos, 2013). If there is sufficient quantitative data, such as time series or exploratory data, statistical quantitative methods such as ANN can be applied. Otherwise, it is appropriate to use qualitative judgmental methods such as the Delphi method if there are insufficient quantitative data available. Additionally, situations other than these two, such as predicting the availability of a new energy source, can be expressed as unpredictable events (Makridakis et al., 1997). The figure showing the characteristics of forecasting methods that are more convenient to use with respect to the available data is as follows.

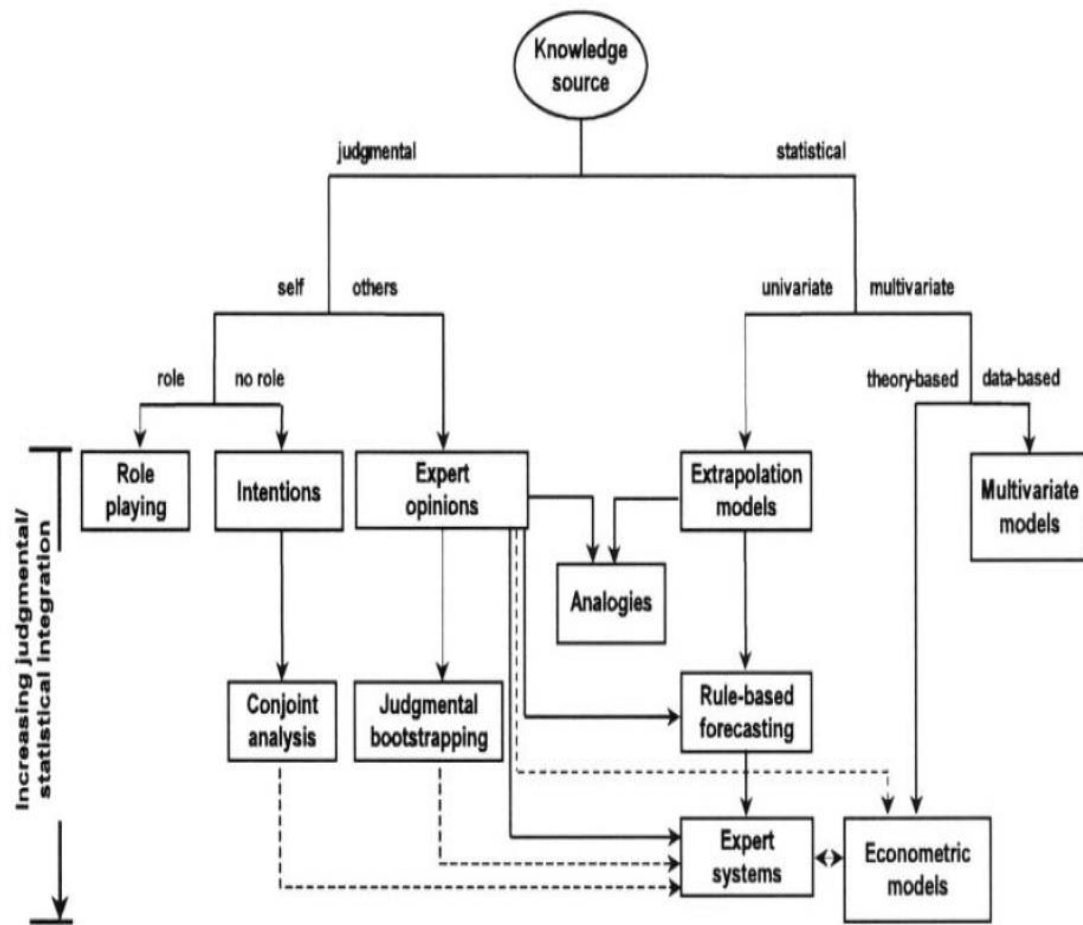


Figure 1.3 Characteristics of Forecasting Methods and Their Relationship (Makridakis et al., 1997)

In summary, it can be said that forecasting has an important place in every department and it is a useful and effective tool for businesses when used appropriately. In the process of choosing the method, the available data should be thoroughly examined, and the method should be chosen according to the data. In this study, the use of ANN, a quantitative method, was deemed appropriate to predict demand.

1.2. Power Transformers and Transformer Theory

Transformers are devices that change the voltage and current to transfer power by electromagnetic induction between circuits via their windings and magnetic core. From household appliances to industrial appliances, transformers used in various applications (Winders, 2002). They are static devices that have the ability to transfer electrical energy from one circuit to another, and while doing this, they did not cause frequency changes between circuits and played an important role in the widespread use of AC type electrical systems we use today. (Kulkarni & Khaparde, 2013)

Before the invention of transformers, that is, at the beginning of the nineteenth century, the first power generation facilities and companies had to use electrical systems working with the DC principle because they could not benefit from those features of transformers, as mentioned above. After the invention of power transformers, pioneering companies that produce electricity have seen how important potential and capability transformers have in electricity generation, distribution and transmission, and have made electricity generation more efficient by mounting these devices to their electrical systems. (Heathcote, 1998)

Power transformers are divided into two categories: step-up and step-down transformers. Step-up power transformers transmit the generated electricity which is produced with AC to high voltages with low current to receiving stations which are placed far from where electricity is produced but close to cities. Following this, step-down transformers reduce the voltage and make the electrical energy ready for use in homes or other places where electricity is needed, together with distribution centers. As a consequence of increased voltage, electrical energy can be transferred farther away from places where electricity is produced and utilized more conveniently. (Kulkarni & Khaparde, 2013)

The following figure illustrates the basic power generation, transfer, and distribution system, along with the utilization of the power transformers used in different phases of this power system, where the step-up and step-down transformers change the voltage and current level of electricity at distinct levels.

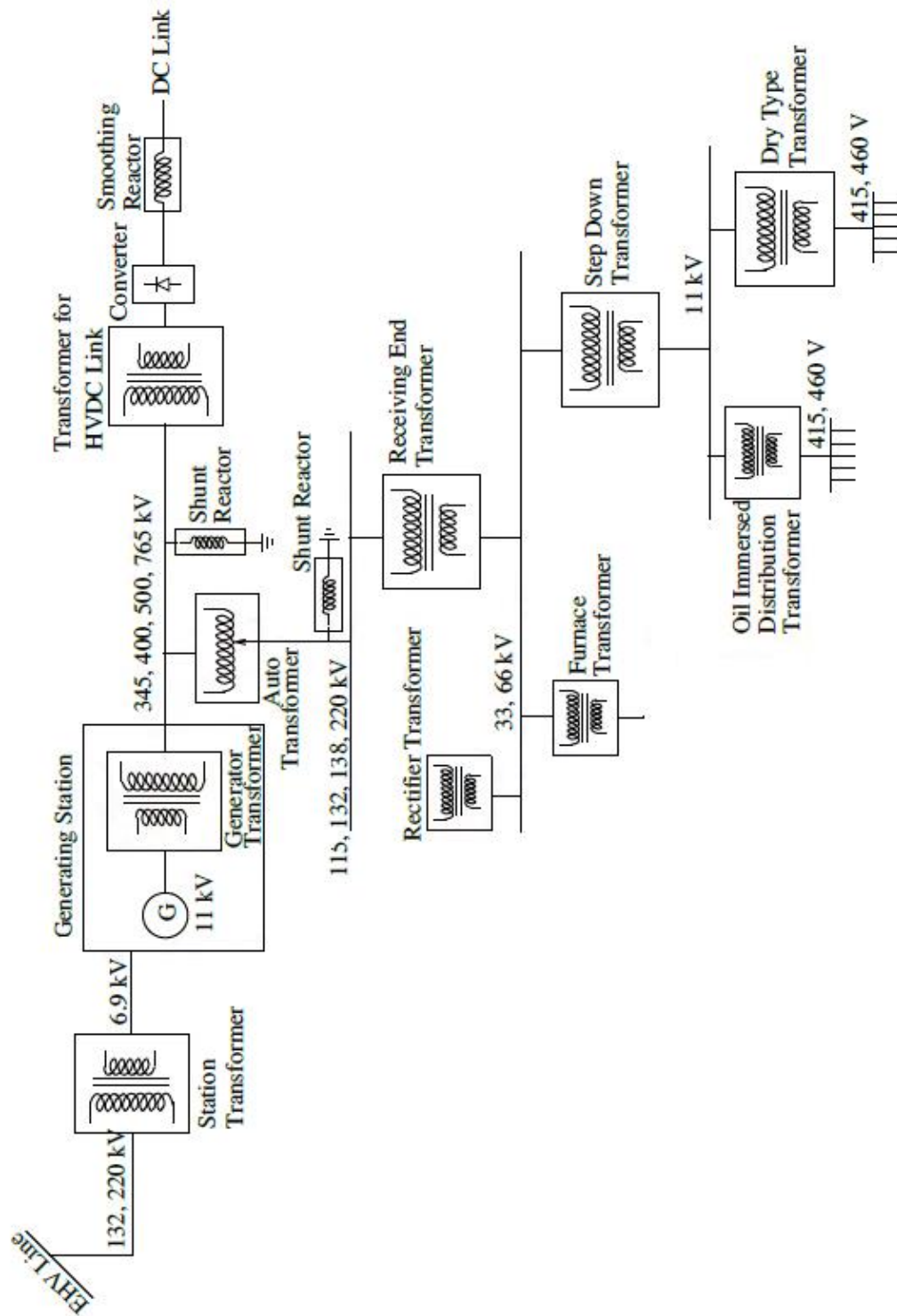


Figure 1.4 Utilization of the power transformers in different phases (Kulkarni & Khaparde, 2013)

Power transformers consist of two or more windings (coils) combined with a magnetic field (Alonzo, 2010) and are mostly grouped according to their power handling capacity (kVA), coolant type, or substation type (Fernandez et al., 2013). Oil immersed hermetically sealed transformers, oil immersed medium power transformers, oil immersed distribution transformers, oil immersed transformers with conservator, cast resin transformers and contemporary ecofriendly green power transformers could be given as an example of this type transformers. In this research only the power transformers which have been widely used by power plants, electricity utility companies and industrial facilities, have been taken into consideration.

The principle of electromagnetic induction forms the fundamental of power transformers. Once a voltage is applied to one of the windings of the power transformer, a current creates an alternating m.m.f. flow, together with a flux in the core. This flux interconnects the windings and induces an e.m.f in each of the windings. In the primary winding this is called back-e.m.f and in theory the transformer will resist the primary applied voltage to the amount that no current flows. However, in reality, there is no electrical connection between these two circuits, but a magnetic connection formed by induction. When alternating voltage is applied to the primary winding for the transformer to work, this voltage will cause current to flow through the primary winding. This current flowing through the core creates a flux that completes its circuit, which induces the E2 voltage in the secondary winding. If there is a load on the secondary side, the current and voltage generated by the induced E2 voltage are applied to the secondary winding and the load (Heathcote, 1998). Thus, the following relationship emerges;

$$E_1/E_2 = N_1/N_2 \quad (1.1)$$

Hence;

$$I_1 \cdot N_1 = I_2 \cdot N_2 \quad (1.2)$$

where E, I, and N are the induced voltage, current, and number of turns, respectively. Additionally, Faraday and Lenz states that the magnitude of the induced voltage is proportional to the rate of change of the flux linkage and the magnetic field created by the current in that

conductor as a result of the induced voltage creates a magnetic field in the opposite direction to the magnetic field that induces the voltage in the same conductor respectively. By taking these discourses as a reference the following formula could be revised accordingly; (Heathcote,1998)

$$E/N=K \cdot \Phi_m \cdot f \quad (1.3)$$

where K , Φ_m , and f are constant, the maximum value of the total flux linking that turns (in Webers), and supply frequency (Hz), respectively. Bearing in mind the above information a theoretical power transformer could be presented with the visual in the following figure.

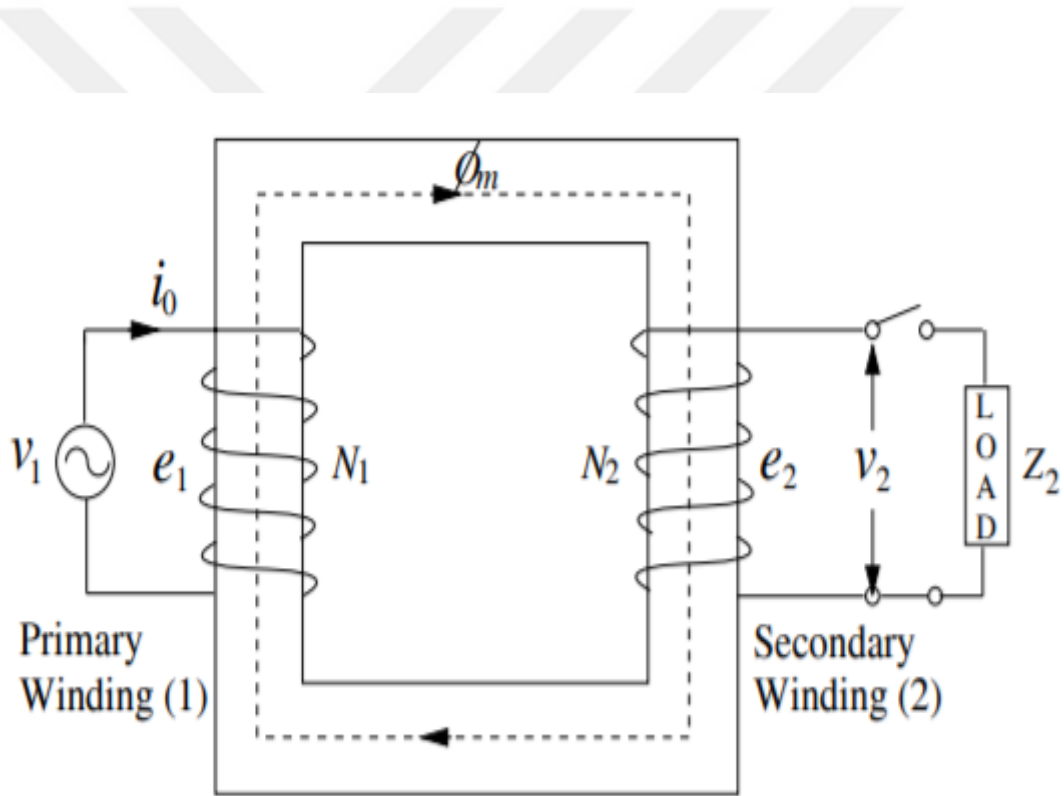


Figure 1.5 Transformer in no-load condition. (Kulkarni & Khaparde, 2013)

1.3. Transformer Market

With the increasing population over time, the need for new energy facilities is increasing (Fofana, et al., 2002). Each energy system required generators, interconnectors, and distribution points. To fulfill this requirement and to make the energy usable, voltage levels must be changed via power transformers (Fernandez, et al., 2013). This possibility provided by power transformers makes power transformers an indispensable part of electrical systems (Bashi et al., 2006).

The machinery industry is a highly important sector in all developed countries and defined as a prior sector. Withal, the majority of production in the power machinery sector is in semi-finished and investment products. In Turkey, approximately 8,400 manufacturer operates in the power machine sector employing around 186,987 people (Sanayi Genel Müdürlüğü, 2020). The most important part of the demand of the sector is electrical equipment, electric motors, generators and transformers, and electricity distribution and control devices. (TR Ministry of Treasury and Finance, 2016; Gündüz, 2016)

Power transformers, distribution transformers, and measurement transformers are widely produced in Turkey, and the production of power and distribution transformers is carried out by nine large and medium-sized companies. The capacity of these facilities is around 27,000 MVA and the engineering quality of power transformers is equivalent to EU countries. In addition, 12 large companies are capable of producing transformers up to 400 KV (TR Ministry of Treasury and Finance, 2016).

Turkey's exports of electrical machinery and equipment in 2020 amounted to approximately 9.29 billion dollars (International Trade Center, 2020). Insulated wires and cables, electrical household devices, monitors and projectors, power transformers, electrical circuit equipment, electrical control distribution boards/panels, generators, and accumulators are important groups in export as. In terms of exports, it is estimated that the structure consisting of insulated wires and cables, electrical household appliances, and transformers, which have a significant share in the sector's total, will continue (TR Ministry of Treasury and Finance, 2016).

Table 1-1 List of products exported by Turkey (Product:85)

Code	Product label	Exported value in 2018 (US Dollar thousand)	Exported value in 2019 (US Dollar thousand)	Exported value in 2020 (US Dollar thousand)
8544	Insulated "incl. enameled or anodized" wire, cable "incl. coaxial cable" and other insulated . . .	2,142,588	2,072,181	2,210,014
8516	Electric instantaneous or storage water heaters and immersion heaters; electric space-heating . . .	1,047,813	1,075,180	1,218,559
8528	Monitors and projectors, not incorporating television reception apparatus; reception apparatus . . .	1,489,484	1,342,029	1,135,430
8504	Electrical transformers, static converters, e.g. rectifiers, and inductors; parts thereof	736,488	793,463	933,262
8536	Electrical apparatus for switching or protecting electrical circuits, or for making connections . . .	495,419	505,081	566,992
8537	Boards, panels, consoles, desks, cabinets, and other bases, equipped with two or more apparatus . . .	490,508	498,459	514,603
8507	Electric accumulators, incl. separators therefore, whether or not square or rectangular; parts . . .	386,687	398,842	359,460
8503	Parts suitable for use solely or principally with electric motors and generators, electric . . .	172,064	255,332	321,995
8502	Electric generating sets and rotary converters	190,897	226,951	288,654
8512	Electrical lighting or signaling equipment (excluding lamps of heading 8539), windscreen wipers, . . .	280,988	270,455	246,701
8511	Electrical ignition or starting equipment of a kind used for spark-ignition or compression-ignition . . .	137,126	124,819	224,529
8501	Electric motors and generators (excluding generating sets)	160,649	192,730	221,961
8541	Diodes, transistors, and similar semiconductor devices; photosensitive semiconductor devices, . . .	24,604	36,546	171,174
8538	Parts suitable for use solely or principally with the apparatus of heading 8535, 8536, or 8537, . . .	145,083	136,032	143,634
8517	Telephone sets, incl. telephones for cellular networks or other wireless networks; other . . .	195,254	216,088	116,282
8535	Electrical apparatus for switching or protecting electrical circuits, or for making connections . . .	105,271	80,788	88,953
8543	Electrical machines and apparatus, having individual functions, n.e.s. in chapter 85 and parts . . .	50,120	52,290	66,319
8515	Electric, incl. electrically heated gas, laser or other light or photon beam, ultrasonic, electron . . .	49,641	57,933	60,068
8508	Vacuum cleaners, incl. dry cleaners, and wet vacuum cleaners	42,253	34,150	53,264
8514	Industrial or laboratory electric furnaces and ovens, incl. those functioning by induction . . .	49,772	53,886	45,264

From the previous table, it is possible to see that the majority of these exports are made to EU countries, and the most exported countries in these product types are the United Kingdom, Germany, France, Iraq, and Italy, respectively from the following illustration.



Figure 1.6 List of importing markets for a product exported by Turkey part (Product:85)

As seen in the figure above, while there is a decrease in exports in some countries, an increase is noteworthy in some countries. For instance, an increasing trend is observable in trade and export amount with the United Kingdom, Germany, and Italy with approximately 1.2, 1.0 and .4 billion dollars respectively, while a decreasing trend in trade and export amount with France is noticeable with the amount around .55 billion dollars in product group which is only including the product family consist of insulated wires and cables, electrical household devices, monitors and projectors, power transformers, electrical circuit equipment, electrical control distribution boards/panels, generators, and accumulators. As can be seen in the later stages of the research, only the top five countries with the highest export volume will be evaluated, as they constitute approximately 33 percent of Turkey's total number of exports in this product group.

When world trade data were analyzed, the global electric and electronics industry reached approximately \$ 4 trillion in 2019. This amount corresponds to approximately 5% of the global economy, which is roughly \$ 80 trillion on a nominal scale. (Sanayi Genel Müdürlüğü, 2020).

Table 1-2 List of exported products for Product 85

Product code	Product label	Exported value in 2018, US Dollar thousand	Exported value in 2019, US Dollar thousand
8542	Electronic integrated circuits; parts thereof	704,073,746	702,314,556
8517	Telephone sets, incl. telephones for cellular networks or other wireless networks; other ...	578,973,876	566,076,066
8544	Insulated "incl. enameled or anodized" wire, cable "incl. coaxial cable" and other insulated ...	131,714,962	127,322,172
8541	Diodes, transistors, and similar semiconductor devices; photosensitive semiconductor devices, ...	112,494,629	114,228,216
8536	Electrical apparatus for switching or protecting electrical circuits, or for making connections ...	108,939,911	104,798,359
8504	Electrical transformers, static converters, e.g. rectifiers, and inductors; parts thereof	100,308,416	100,004,826
8528	Monitors and projectors, not incorporating television reception apparatus; reception apparatus ...	94,410,918	90,682,178
8537	Boards, panels, consoles, desks, cabinets, and other bases, equipped with two or more apparatus ...	68,796,594	68,898,991
8507	Electric accumulators, incl. separators therefore, whether or not square or rectangular; parts ...	57,029,127	61,875,090
8529	Parts suitable for use solely or principally with transmission and reception apparatus for ...	57,621,479	58,731,767
8501	Electric motors and generators (excluding generating sets)	56,077,601	55,492,479

Considering the data in the above table, it is possible to state that the top five product families exported worldwide are electronic integrated circuits, telephone sets, insulated wires and cables, diodes, transistors, and similar semiconductor devices, electrical apparatus for switching or protecting electrical circuits, and for making connections and electrical transformers and static converters, respectively. To sum up, the product group (Product 85) that the power transformers are in, is a group that is highly demanded, exported, imported, and widely traded both by Turkey and global scale and the numbers solidify and the emphasize the importance of the sector in case.

1.4. Factors Affecting Demand for Power Transformers

The factors, 9 in total, affecting the demand for power transformers have been defined thanks to the authorities of STD Power Transformer Company which is one of the most important power transformer production facilities, exporting more than eighty-five percent of its products to 17 different countries and located at Turkey, Adana. Accordingly, the factors affecting the demand for power transformers can be listed as production volume, population by country, total power installed, total power generated, electricity consumption, number of enterprises, electricity consumption of the topmost exported countries, population of the topmost exported countries, and number of installed power transformers.

1.4.1. Production Volume

It is predicted that the market, including power machines such as power transformers, is continuously expanding, and increasing the amount of transformer production will positively affect the demand for the product, as it will positively affect the supply offered to the market. In this context, considering the compatibility of the data obtained, between the periods from 2007 to 2019 the total production amount of liquid dielectric transformers (LDT) and transformers having a power handling capacity of more than 500 kVA have been into consideration and following table have been proposed.

Table 1-3 Annual production volume of PT (TUIK)

Year	LDT and PT in Total (Unit)
2021	54630
2020	48105
2019	37089
2018	37747
2017	29396
2016	21565
2015	19512
2014	36276
2013	34335
2012	62842
2011	39190

2010	48639
2009	35111
2008	36783
2007	32418
2006	38377
2005	28496
2004	35820
2003	35553
2002	35405
2001	35249
2000	35397
1999	35230
1998	35311
1997	35253
1996	35082
1995	34928
1994	34774
1993	34870
1992	34725
1991	34634
1990	34621
1989	34469
1988	34465
1987	34427
1986	34289
1985	34194
1984	34138
1983	34056
1982	33998

In the above table, LDT and PT have more than 500 kVA power-handling capacity, are obtained cumulatively, and in the later stages of the research, the total of these products that belong to

the same product range will be evaluated concerning their summation. However, as the official production amount of PTs is just kept until 2005 remaining data has been generated via linear regression by taking the gross national product into account which refers to the summation of total services and manufactured units in a certain currency in basic and as the official sources of Turkey have changed to data collection method for GNP in period 2006 the following data has been sourced from outbound (The World Bank, 2022) sources.

1.4.2. Country Population

It is believed that population data have a direct and indirect positive effect on electricity usage. However, in terms of the compatibility of the data obtained previously, the population data were only based on the number of citizens who had an official residence in the country until 2021. The following table illustrates the population of Turkey based on the years considering mentioned issues.

Table 1-4 Address-Based Population of Turkey (TUIK)

Year	Population
2021	84680273
2020	83614362
2019	83154997
2018	82003882
2017	80810525
2016	79814871
2015	78741053
2014	77695904
2013	76667864
2012	75627384
2011	74724269
2010	73722988
2009	72561312
2008	71517100
2007	70586256
2006	68626337
2005	67743052

2004	66845635
2003	65938265
2002	65022300
2001	64100297
2000	63174483
1999	62243779
1998	61308204
1997	60302413
1996	59442502
1995	58522320
1994	57613441
1993	56713073
1992	55815175
1991	54911233
1990	53994605
1989	53065801
1988	52125597
1987	51168270
1986	50186914
1985	49178079
1984	48138191
1983	47072603
1982	45997940

1.4.3. Total Power Installed

Total installed power is an important indicator reflecting the number of electricity generation facilities in the country. It is thought that this indicator may increase with the improvement work to be carried out in existing electricity generation facilities, as well as with newly established power plants. As stated before, improvements, upgrades in already existing energy systems or a new power plant establishment, which may include hydroelectric, nuclear, solar, coal, gas or diesel-fired, geothermal, tidal or wind power plant, requires power transformers in different components and levels of these energy systems such as generation, interconnection or

distribution, to make the electricity useable for both home or industrial use. For these reasons, the total installed power statistics have been compiled and the table below has been created to illustrate the installed power stats which is covering the data from 1982 to 2021.

Table 1-5 Total Power Installed by Power Plants (TEİAŞ)

Year	Total Power Installed (MW)
2021	99819.6
2020	95890.6
2019	91267.0
2018	88550.8
2017	85200.0
2016	78497.4
2015	73146.7
2014	69519.8
2013	64007.5
2012	57059.4
2011	52911.1
2010	49524.1
2009	44761.2
2008	41817.2
2007	40835.7
2006	40564.8
2005	38843.5
2004	36824.0
2003	35587.0
2002	31845.8
2001	28332.4
2000	27264.1
1999	26119.3
1998	23354.0
1997	21891.9
1996	21249.4
1995	20954.3

1994	20859.8
1993	20337.6
1992	18716.1
1991	17209.1
1990	16317.6
1989	15808.2
1988	14520.6
1987	12495.1
1986	10115.2
1985	9121.60
1984	8461.60
1983	6935.10
1982	6638.60

1.4.4. Total Power Generated

Power and energy are different measuring units from each other. Briefly, power is the immediate proportion at which energy is set free. While power is expressed in watts (W) or kilowatts (kW), energy is expressed in kilowatt-hours (kWh) or gigawatt hours (GWh) owing to the relationship between power and time (Plett, 2015). For this reason, gross energy generation and the total power installed decomposes, and due to the potential of the total generated energy offered to the market has the potential of increasing demand for the power transformers following table includes gross generation amount starting from 1982 to 2021 have been constructed.

Table 1-6 Gross Generation of Energy (TEİAŞ)

Year	Gross Generation (GWh)
2021	334723.1
2020	306703.1
2019	303897.6
2018	304801.9
2017	297277.5
2016	274407.7

2015	261783.3
2014	251962.8
2013	240154.0
2012	239496.8
2011	229395.1
2010	211207.7
2009	194812.9
2008	198418.0
2007	191558.1
2006	176299.8
2005	161956.2
2004	150698.3
2003	140580.5
2002	129399.5
2001	122724.7
2000	124921.6
1999	116439.9
1998	111022.4
1997	103295.8
1996	94861.70
1995	86247.40
1994	78321.70
1993	73807.50
1992	67216.80
1991	60499.30
1990	56811.70
1989	52601.70
1988	48430.00
1987	44925.00
1986	40471.40
1985	36361.30
1984	33266.50

1983	29567.60
1982	28324.90

1.4.5. Consumption of Energy

In this study, it is predicted that the fluctuations in energy consumption may affect the demand for power transformers positively or negatively. It is anticipated that the increase in energy consumption may pave the way for new energy facilities or facility improvements. Therefore, as stated before, since the new facilities and improvements will require new transformers, the energy consumption indicator has found a place to itself in this study. As in the previous data, the energy consumption data was taken from 1982 to 2021 in terms of compatibility with other data, and the table below is constructed on this amendment. However, since the official data provider institution in Turkey did not disclose the data kept for some years as a result of confidentiality or impossibility, the linear regression method was applied in this parameter as well. The official data provider has disclosed data for the years 2021 to 2007 and data for the years 2000 and 1990. The remaining data were obtained from the own resources of the official electricity distribution and transmission company in Turkey. While applying linear regression to find the data that is not available, the average of the maximum and minimum daily electricity consumption in Turkey between 1982 and 2021 have been utilized. (TEİAŞ, 2022)

Table 1-7 Net Energy Consumption (TEİAŞ)

Year	Net Energy Consumption (GWh)
2021	286691
2020	262702
2019	257273
2018	258232
2017	249023
2016	231204
2015	217312
2014	207375
2013	198045
2012	194923

2011	186100
2010	172051
2009	156894
2008	161948
2007	155135
2006	140419
2005	132435
2004	120149
2003	115769
2002	113551
2001	105667
2000	98296
1999	98925
1998	86368
1997	84286
1996	71670
1995	65427
1994	61217
1993	53853
1992	32771
1991	42439
1990	46820
1989	38099
1988	34351
1987	32781
1986	29244
1985	24931
1984	23879
1983	19490
1982	19592

1.4.6. Number of Enterprises

In this study, the following table, which involves the number of enterprises from 1982 to 2021, was constructed considering that the enterprises operating in the manufacturing sector, which include large-scale enterprises and SMEs, are effective in increasing the demand for power transformers. Enterprises including wholesale, retail trade, hotels, restaurants, gas-water suppliers, health, social work, transportation, broadcasting activities, finance, insurance, storage, communication, real estate, education and other community or personal service activities have been excluded since the data collected includes businesses with multiple inclusions. Therewithal, the number of enterprises operating in electricity generation have also been excluded since in previous steps another indicator related with electricity generation had already been gathered. In addition, neither the official data provider institute in Turkey nor the Chamber of Commerce published the data before 2003. The missing data have similarly been calculated using linear regression based on the population of the country.

Table 1-8 Number of Enterprises (TÜİK)

Year	Number of Enterprises
2021	3398029
2020	3304088
2019	3228421
2018	3160371
2017	3100412
2016	2981381
2015	2941233
2014	2888180
2013	2847725
2012	2800060
2011	2737278
2010	2678787
2009	2631085
2008	2583099
2007	2567704
2006	2473841

2005	2393578
2004	2002834
2003	1740353
2002	2047461
2001	1984856
2000	1921992
1999	1858796
1998	1795269
1997	1726975
1996	1668585
1995	1606104
1994	1544390
1993	1483253
1992	1422285
1991	1360906
1990	1235599
1989	1164046
1988	1091093
1987	1016205
1986	939070
1985	859392
1984	777491
1983	694552
1982	686977

1.4.7. Electricity Consumption of The Topmost Exported Countries

As underlined, power transformers are highly traded products worldwide, including Turkey. Turkey's exports of electrical and electronic products are approximately \$11.5 billion, of which \$6.6 billion is made with EU countries. Among the EU countries, the top five countries to which Turkey export the most in this sector are the United Kingdom, Germany, France, Italy and Spain, respectively. However, exports among others countries higher than \$200 million could be sorted as Iraq, Israel, United Arab United, Saudi Arabia and the USA respectively. (TİM,

2019). In this research, when evaluated together with the other countries outside the European Union, the most exported top five countries with the product family including power transformers are evaluated as United Kingdom, Germany, France, Iraq and Italy, respectively as can be seen in Figure 1. Considering all these, it is concluded that the demand may not only be national or regional, but may also could be affected from countries with whom the Turkey trade mostly. In this context, since the increase in energy consumption of topmost traded countries mentioned above may pave the way for the new energy facilities or facility improvements and indirectly the demand of power transformer the electricity consumption of the five most exported countries has been drawn up from 1982 to 2021 as seen below.

Table 1-9 Electricity Consumption of The Topmost Exported Countries

Year	UK (GWh)	Germany (GWh)	France (GWh)	Iraq (GWh)	Italy (GWh)
2021	286050	511590	447622	39299	298286
2020	280601	499557	425566	35726	287784
2019	295475	517805	449423	43044	284679
2018	300661	512934	449634	43051	306881
2017	299873	518953	455183	39992	305383
2016	304276	517554	457721	38733	295508
2015	303891	514948	450813	35585	297180
2014	303347	511757	441558	34313	291084
2013	316844	522310	463881	45042	297288
2012	318757	524524	462296	35076	307220
2011	318253	524877	451005	38841	313792
2010	329258	532006	474020	35191	309885
2009	321748	497471	446831	25007	299915
2008	341822	527755	462763	21839	319037
2007	342263	529527	450009	21828	318953
2006	349352	552295	447833	25225	317533
2005	339575	546707	452118	20997	309817
2004	336872	550820	446957	29681	304490
2003	334056	544414	435439	24995	299789
2002	333889	532517	418210	29531	290960

2001	330598	520195	420453	28131	285492
2000	324025	509081	409744	27277	279320
1999	316944	497056	401221	25568	267284
1998	312449	497445	393359	26781	260883
1997	310576	492931	381735	26667	253485
1996	295855	489296	382650	26080	246033
1995	285311	482523	367839	26180	243676
1994	286133	475055	363556	25217	236957
1993	281472	477208	356080	23598	229536
1992	281050	485685	355161	22640	229821
1991	274432	491098	345386	17821	224429
1990	270402	497855	323043	19520	219007
1989	265349	381385	316531	25996	211842
1988	258891	374809	309385	24601	186849
1987	250314	371555	301429	21657	193549
1986	241666	370889	292670	19875	183551
1985	234099	370071	277937	18716	178066
1984	224321	370267	260163	16334	174155
1983	219466	372023	247079	14449	164854
1982	216011	373657	242635	12186	165723

As seen in the table, while the annual electricity consumption of some countries increases, that of some countries shows a decreasing trend. As stated before, the data were analyzed in terms of consistency between 1982 and 2021, and in order to ensure this, the data that were not available in some years were compiled from different sources that were consistent with each other. For instance, the electricity consumption data of Iraq between 2007 and 2018 have been gathered from the International Energy Agency (IEA), while the electricity consumption data for 2019-2021 and 1982-2007 of the same country have taken from Enerdata. However, since the pre-reunification data of Germany in 1990 were not available, the data from this date to 1982 were calculated with linear regression similarly and the population of Germany was taken into account during this calculation. Furthermore, the five countries mentioned above will not be evaluated separately in this research as it does not match the theme of the study and the total

annual electricity consumption of these topmost exported countries will be considered as an indicator as it could be seen with the parameters given below.

Table 1-10 Total Electricity Consumption of Topmost Exported Countries

Year	Total Electricity Consumption of Topmost Exported Countries (GWh)
2021	1582847
2020	1529234
2019	1590426
2018	1613161
2017	1619384
2016	1613792
2015	1602417
2014	1582058
2013	1645364
2012	1647873
2011	1646768
2010	1680360
2009	1590973
2008	1673216
2007	1662579
2006	1692238
2005	1669214
2004	1668820
2003	1638693
2002	1605107
2001	1584869
2000	1549446
1999	1508073
1998	1490917
1997	1465394
1996	1439914
1995	1405529

1994	1386918
1993	1367894
1992	1374357
1991	1353166
1990	1329827
1989	1201103
1988	1154535
1987	1138505
1986	1108652
1985	1078889
1984	1045240
1983	1017871
1982	1010212

1.4.8. Population of The Topmost Exported Countries

As stated in the previous sections, it can be said that the export and import potential of PTs is quite high, as can be seen with the numbers given previously. This situation shows that the total demand for the product is not only regional, but also international. Moreover, this situation emerges the requirement of gathering the data collected on a global scale in order to progress in a more solid and reliable way. In this context, such as obtaining annual electricity consumption data of the topmost exported countries, the population data of these countries are also have been gathered since it considered as a significant indicator in previous steps that directly affects electricity consumption and demand for power transformers and following table have been constructed which is covering the population of topmost traded countries United Kingdom, Germany, France, Iraq and Italy respectively starting from 2007 to 2019.

Table 1-11 Populations of The Topmost Exported Countries

Year	UK	Germany	France	Iraq	Italy
2021	68709000	83237124	65447454	41563520	59066225
2020	67081000	83155031	65284389	40222493	59449527
2019	66796800	83166711	65096768	39309783	60297396
2018	66435600	83019213	64844037	38433600	60421760
2017	66040200	82792351	64639133	37552781	60536709
2016	65648100	82521653	64468792	36610632	60627498
2015	65110000	82175684	64300821	35572261	60730582
2014	64596800	81197537	64027958	34411951	60789140
2013	64105700	80767463	63697865	33157050	60233948
2012	63705000	80523746	63375971	31890011	59539717
2011	63285100	80327900	63070344	30725300	59379449
2010	62759500	81751602	62765235	29741976	59277417
2009	62260500	81802257	62465709	28973162	59095365
2008	61823800	82002356	62134866	28385746	58826731
2007	61319100	82217837	61795238	27911248	58438310
2006	60827100	82314906	61399733	28905607	58143979
2005	60413300	82437995	60963264	28698684	57969484
2004	59950400	82500849	60505421	27858948	57685327
2003	59636700	82531671	60101841	27068823	57313203
2002	59365700	82536680	59685899	26255343	57059007
2001	59113000	82440309	59266572	25425663	56974100
2000	58886100	82259540	58858198	24628858	56942108
1999	58684400	82163475	58496613	23820734	56916317
1998	58474900	82037011	58298962	23053488	56906744
1997	58314200	82057379	58116018	22330770	56890372
1996	58164400	82012162	57935959	21639332	56860281
1995	58024800	81817499	57752535	20948858	56844303
1994	57862100	81538603	57565008	20248073	56843400
1993	57713900	81338093	57369161	19295818	56831821
1992	57584500	80974632	57110533	18385673	56797087

1991	57438700	80274564	56840661	17846378	56758521
1990	57237500	79753227	56577000	17658381	56719240
1989	57076500	62679035	56269810	17442590	56671781
1988	56916400	61715103	55966142	16987346	56629288
1987	56804000	61238079	55681780	16598019	56601931
1986	56683800	61140461	55411238	16148221	56596155
1985	56554000	61020474	55157303	15753856	56593071
1984	56409300	61049256	54894854	15377032	56576718
1983	56315700	61306669	54649984	14903737	56564074
1982	56290700	61546101	54335000	14467682	56543548

As in the previous section, the populations of the top five countries with the highest export numbers will not be evaluated separately, and the study will progress on the total population of these countries. Within this framework, the total populations of these countries covering 1982 and 2021 were calculated, taking into account the compatibility of the data as before, and the following table was constructed:

Table 1-12 Total Population of The Topmost Exported Countries

Year	Total Population of The Topmost Exported Countries
2021	318023323
2020	315192440
2019	314667458
2018	313154210
2017	311561174
2016	309876675
2015	307889348
2014	305023386
2013	301962026
2012	299034445
2011	296788093
2010	296295730

2009	294596993
2008	293173499
2007	291681733
2006	291591325
2005	290482727
2004	288500945
2003	286652238
2002	284902629
2001	283219644
2000	281574804
1999	280081539
1998	278771105
1997	277708739
1996	276612134
1995	275387995
1994	274057184
1993	272548793
1992	270852425
1991	269158824
1990	267945348
1989	250139716
1988	248214279
1987	246923809
1986	245979875
1985	245078704
1984	244307160
1983	243740164
1982	243183031

1.4.9. Number of Installed Power Transformers

Another important indicator, namely, the number of installed power transformers, has been defined to understand the demand for power transformers. This factor includes the data of annual development of transmission transformers and directly correlated with the demand of power transformers in general. However, power transformers are not only utilized for transmission duty but also other components of electricity systems such as distribution and generation. It should be underlined that the following data held by official electricity supplier company is only includes the number of installed power transformers in transmission mission. Although it reflects a limited part of the utilization area, considering that these data can directly affect the demand positively or negatively, it was deemed appropriate to include it in the study.

Table 1-13 Number of Installed Power Transformers

Year	Number of Installed PTs
2021	2083
2020	1966
2019	1889
2018	1826
2017	1750
2016	1675
2015	1628
2014	1550
2013	1505
2012	1425
2011	1357
2010	1317
2009	1272
2008	1241
2007	1173
2006	1130
2005	1096
2004	1089

2003	1072
2002	966
2001	1003
2000	978
1999	965
1998	926
1997	882
1996	834
1995	811
1994	788
1993	764
1992	756
1991	712
1990	676
1989	651
1988	628
1987	609
1986	591
1985	644
1984	633
1983	604
1982	588

1.5. Demand for Power Transformers

It would not be wrong to state that sales in manufacturing systems are synonymous with demand. In this context, the historical sales data of power transformers between 1982 and 2021 were utilized to forecast the demand for the same product. The data have been obtained from the official electricity supplier company and have been limited between 2007 and 2021 by considering the compatibility with other data utilized in study same approach on chapter 1.4.1. have been applied.

Table 1-14 Historical Sales Amount of Power Transformers

Year	Sales (unit)
2021	53984
2020	46718
2019	36913
2018	36375
2017	27533
2016	22307
2015	19934
2014	35147
2013	34862
2012	59520
2011	37252
2010	48233
2009	35115
2008	34469
2007	32455
2006	36639
2005	36530
2004	36401
2003	36273
2002	36202
2001	36127
2000	36198
1999	36118
1998	36157
1997	36129
1996	36047
1995	35973
1994	35899
1993	35945
1992	35875

1991	35832
1990	35825
1989	35753
1988	35750
1987	35733
1986	35666
1985	35620
1984	35594
1983	35554
1982	35526

2. LITERATURE REVIEW

The significance of forecasting, the decisiveness and effectiveness of the role it plays within the businesses, together with the convenience and foresight it emerged, have been mentioned earlier on previous pages. In this context, it is reasonable to say that many scientific studies have been carried out in order to predict the future much more prominent in recent years through the agency of statistical techniques (1994). In this framework, it has been planned to conduct a literature review in this section in order to both construct a more sophisticated and solid hypothesis mesh and to unfold the previous studies about the principal object. Furthermore, the review planned to be made, had been carried out in three main headings as demand forecasting, ANNs in forecasting applications and the as the last one as a combination of previous two subheadings, demand forecasting with ANNs. Finally, the chapter will be concluded with the presentation of the hypothesis which takes the literature review into the account.

2.1. Demand Forecasting Applications from Various Topics

Tsao et al. (2022) proposed a demand forecasting model and applied it to a b2b server company located in the United States. They initiated their model with collecting historical sales data and applied machine learning and time series algorithms. Following that they improved the model with information gathered from various sources such as Google Trends and found out that their improved model has outperformed the model that they constructed previously dramatically. Moreover, Hu et al. (2022) have constructed a demand forecasting model in purpose of forecast the tourism demand in Hong Kong. They took online review data that produced by tourists such as shopping, attraction points, hotels, etc. in to account. They indicated that mixed data sampling models was the most accurate model when high volume online review available.

Rožanec et al. (2021) investigated and compared various forecasting approaches to determine the future demand of OEM companies in Europe. They have utilized various data sources. They concluded that support vector machine, random forest regressor and voting ensemble algorithms were the most suitable algorithms that trained on data of same type demand. They also indicated that machine learning models have performed significantly when they grouped the products regarding their demand patterns or historical demand. Huang et al. (2021) have

created a demand forecasting model to forecast the regional logistics demand in a perspective that focuses e-commerce indicators. They found out that the back propagation algorithm they utilized is capable of producing accurate results for regional logistics demand.

Souza et al. (2020) evaluated four different fuzzy inference system-based forecasting models on twenty-seven different nail polish products of multinational beauty companies, and compared the results obtained with the firm's own forecasts, which were mostly determined by qualitative forecasting techniques. They noted that fuzzy inference system-based forecasting models is more efficient and preferable in long-term demand forecasting, and quantitative forecasting techniques currently used by the firm is also preferable in the in short-term demand forecasting.

IHME and Murray (2020) in his study in America for the covid 19 epidemic that the world suffered from, forecasted the number of patients who may need ventilator and intensive care unit and deaths for four months later. Khaldi et al. (2017) proposed a study that endeavors to forecast the demand for blood using two different forecast models, ARIMA and ANN. They evaluated red blood cells, platelets, and plasmas separately and concluded that the model constructed with artificial neural networks produced more consistent predictions than the model created with ARIMA.

Priyadarshi et al. (2019) have executed a number of different forecasting methods such as autoregressive integrated moving average model, support vector regression, long short-term memory, gradient boosting regression and extreme gradient boosting regression and random forest regression in order to find out the most appropriate forecasting model at the retail stage for vegetables and concluded that long short-term memory model and support vector regression model outperformed in comparison with the other methods. Similarly, Yang and Sutrisno (2018) applied an ANN forecasting technique to determine daily demand for goods. They utilized approximately ten million sales date which acquired from bakery franchise operating in China and concluded that number of sales in the beginning of day could be used in order to forecast the rest daily demand.

Sen and Chaudhuri (2017) conducted a study that examined the Indian health sector through six years of time series data to forecast the demand that the health sector may face. They applied six different forecasting methods such as; Holt Winters, auto regression (AR), moving average (MA), ARIMA, partial auto correlation function (PACF) and auto correlation function. They have executed the study in a comparative basis and reached the best forecasting result using ARIMA model and the worst estimation result with Holt Winters model respectively. Hubert et al. (2017) have applied a multivariant ARIMA model in order to forecast the daily demand for industrialized bakery chain found out that it is possible to increase the availability of products and stated the and the model they proposed is effective at the operational level. Tirkes et al. (2017) have discussed trend analysis, decomposition and holt-winters estimation models in a comparative way. They applied these forecasting methods based on the annual monthly sales data of jam and sherbet products and stated that according to the MAPE results they have obtained, the more efficient methods were holt-winters and decomposition models.

Panapakidis and Dagoumas (2016) developed a hybrid model that comprises the wavelet transform, genetic algorithm, adaptive neuro-fuzzy inference system, and feed-forward neural network to examine the fitness of the proposed model and forecast the day-ahead natural gas demand in Greece and obtained satisfactory results. In parallel, Barbosa et al. (2015) have applied Holt-Winters method with past sales data in order to forecast the demand of a medium sized food company which is producing sausages and pastas and the model they suggested reduces the forecast error about five percent with respect to MAPE.

Melikoglu (2013) suggested a linear and logistic equation for medium- and short-term natural gas demand in Turkey, respectively, using population change, demographics, purchasing power parity, and gross domestic product as parameters, and obtained results that are equivalent to the foresight of countries' official petroleum pipeline cooperation. Using short-term half-hourly electricity consumption data Taylor (2013) upgraded standard Holt-Winters method by adjusting an auto-regressive model to the previous model and stated new model is applicable and more accurate for forecasting time-series have more than one seasonal change.

Aksoy et al. (2012) applied an adaptive network-based fuzzy inference system model to forecast the clothing industry. They noted that the input and output criteria of the methodology were

defined with respect to clothing manufacturers and the literature, and by using them, they found that an adaptive-network-based fuzzy inference system model-based demand forecasting system has the potential to help clothing manufacturers in terms of accuracy, effectiveness, and simplicity. Sughanti and Samuel (2012). compiled numerous studies from many countries dealing with electricity demand, and claimed that in these demand forecasts for both conventional and renewable energy sources, mostly ARIMA, ANN, Gray, fuzzy logic, SVR, AGO, PSO and genetic algorithm techniques or combinations of these techniques were utilized. In addition, in the studies they examined, they stated that GNP, energy price, gross output and population were the mostly used parameters as the factor affecting the demand.

Polebitski and Walter (2010) on the other hand have proposed a regression-based forecasting model using twelve years of demographic, economic, weather and consumption data collected from a hundred different mesh blocks. In addition, they found that the water demand in the summer is more elastic considering the economic conditions of citizens and the water prices, than in winter.

Alvisi et al. (2007) developed a demand forecasting model for a project using a pattern-based approach and annual, monthly and weekly water demand data. Similarly, Bougadis et al. (2005) determined the demand for short-term municipal water by using air temperature, rainfall, and previous water consumption data. They made forecasts using ANN, regression and time series and found that the demand for the water was mostly correlated with rainfall. In addition, they stated that they obtained the most satisfactory results with the model they established with artificial neural networks.

Chen and Wang (2005) proposed a genetic algorithm with support vector regression (GA-SVR) using the number of tourist arrivals in China comprising 16 years and found that SVR outperforms the other models applied in studies that are back-propagation and ARIMA, considering the MAPE results they obtained. Burger et al. (2001) compared different forecasting models such as moving average, decomposition, naïve, ARIMA, single exponential smoothing, ARIMA, multiple regression, genetic regression and neural networks by using the number of tourist arrivals destined from United States to Durban (South Africa) indicated that

neural networks gave the best results. Goh and Law (2001) have proposed a time series as SARIMA-MARIMA with interventions by using historical tourist arrival data of Hong Kong in order to forecast the tourism demand for upcoming periods. By taking the different factors in to consideration such as Asian financial crisis, visa policies, bird flu epidemic etc. they concluded that the model that they used (SARIMA-ARIMA) performed best in comparison with other eight models that they applied.

It can be concluded from the review that the demand forecasting phenomenon has drawn attention in many sectors, such as energy, water supply, tourism, petroleum and gases, food, healthcare, and various manufacturing sectors, due to the potential of consolidation and reinforcements for businesses.

2.2. Forecasting Applications with Artificial Neural Networks

The accurate prediction potential of artificial neural networks has also been shown in the literature, where Farimani et al. (2022) proposed and compared two forecasting models, ARIMA and ANN, which investigate the future demand for automotive parts. They applied both models to eight different types of products and concluded that the ARIMA model outperformed the ANN model for six products. Kaya et al. (2022) have investigate three forecasting models namely time series decomposition, ARIMA and ANN which forecasts the automobile sales in Turkey. They utilized various indicators such as price, parity, basic economic indicators etc. and found out that exchange rates, expectations of households, production index and sales amount before one month were the most significant factor that affects the sales. They also point out that ANN model outperformed other models with 1.18% MAPE value.

Aktepe et al. (2021) compared MLR, MNR, support vector regression, and ANN models, which forecast the demand for spare parts by construction machinery companies. They utilized past sales data, number of machines sold worldwide, and economic indicators. They found that although ANN gave better results than regression methods, it could not outperform support vector regression. Additionally, they indicated that monthly impact rate which used as a factor has significant effect on performance of the forecast. Guven and Simsir (2020) have investigated two forecasting models namely support vector machines and ANN which forecasts

the demand of apparel retail industry. They have defined the factors which have the potential of affecting the demand as color, gender, special days and weather conditions and created 20 different models covering 9 types of products. They concluded that ANN produced better results than support vector machines in most datasets.

Yucesan et al. (2019) proposed an ANN that attempts to determine the number of patient arrivals to emergency department in Eastern Turkey. On this purpose they have combined ARIMAX and ANN models by using the historical arrival data of a public hospital with factors, which are defined as corelated with emergency service, namely climatic and calendarial variables and found out that the model proposed have generated satisfactory outputs.

Afana et al. (2018) constructed an ANN to forecast car millage per gallon in the city. They have used different indicators namely model, make, type, origin, drivetrain (transmission) and manufacturer suggested retail price, invoice, cylinders, engine size, horsepower, miles per gallon, weight, length and wheelbase alphabetic and numeric factors respectively. They have obtained reasonable predictions with the 97.50% accuracy and found out that the drivetrain is the most effective factor which influence the miles per gallon most.

Dosdogru et al. (2018) improved an ANN which utilizes ten different metaheuristics to increase the performance of stock market price forecasts. They found that based on the performance metrics, ANNs combined with metaheuristic approaches have great potential to produce successful forecasts. Fente and Singh (2018) have developed an ANN which is using different weather parameters from 2007 to 2017 as humidity, temperature, dew point visibility, precipitation, pressure and windspeed which have gathered from national climate data center in order to make weather forecasts. They have used long-short term memory technique in order to train historical data set and claimed that their experimental results have demonstrated satisfying forecast outputs with accurate predictions. Gocken et al. (2018) have proposed a hybrid model which consist of the combination of both genetic algorithm and harmony search in order to forecast weather condition in six different cities in Turkey. They applied combinations of these three algorithms as single ANN, harmony search ANN and genetic algorithm ANN to optimize

the performance of the forecast. They concluded that the proposed models have produced satisfactory results with less than 6% error rate and capable of forecasting weather remarkably.

Constantino et al. (2016) created a neural network by using overnight stays in hotels between January 2013 and December 2014 to forecast the tourism demand of Mozambique. They have used gross domestic product, consumer price index and exchange rates of the most tourist generating countries as their independent indicator and found out that consumer price index, exchange rates of euro, dollar and South Africa rand is the most effective variables along with the gross domestic product rates of Portugal and USA. Their network performed accurately with the MAPE rate of 6.5%.

Ghiassi et al. (2015) designed a neural network model to forecast movie revenue by upgrading an experimental model of other authors, where the star value, genre, special effects, sequel, number of screens, competition, and rating of motion picture association have been selected as variables influencing the revenues most, to a dynamic artificial neural network model. They upgraded the model by utilizing other factors, namely the production budget, pre-released advertisement expenditures, runtime, and seasonality, and concluded that the novel approach produced excellent outputs with an error rating of 94,1%.

Cuhadar et al. (2014) utilized an ANN to forecast the amount of cruise tourism in Izmir, Turkey. They compared different ANN models, such as multi-layer perceptron, radial basis function, and generalized regression, using cruise tourist arrivals at Izmir Cruise Port between 2005 and 2013. They concluded that radial basis function model outperforms the other models in means of forecasting accuracy. Abd Rahman et al. (2013) have constructed and evaluated three different forecasting models by using ARIMA, ANN and fuzzy time series which forecasts the air pollution in Malaysia. They have used air pollution index which comprised of the rates of NO₂, SO₂, CO₂, O₃ and PM₁₀ in the air. The performance of each model evaluated separately with the rate of root mean square error of each model and they concluded that the model constructed with ANN outperformed with its' accurate outputs.

Pacelli et al. (2011) applied an ANN to forecast the exchange rate of Euro/USD. They used more than 40 variables to build a more reliable and effective neural network, such as indexes of Dow Jones, NASDAQ, CAC, Deutscher Aktien, parity of euro against pounds, yens, dollars, and NZ dollars; spot prices of gold, silver, gas, and oil; and number of different financial indicators. They found out that ANN methodology has the potential of forecast exchange rate of euro dollar parity for three days ahead. Adewole et al. (2011) who conducted a study on a similar subject have designed an ANN in purpose of foreign exchange rate. They have collected around eight hundred daily price quotes between the years of 2003 and 2005. Some of these data have been used for training, and others for validation and testing. They concluded that their novel model outperformed against a well-known Hidden Markov foreign exchange rate forecasting model with the scores of 69.9% against 81.2% respectively.

Baboo and Shereef (2010) designed an ANN using a back-propagation technique to forecast weather temperature. They used five indicator variables: wind direction and velocity, atmospheric pressure, temperature, and relative humidity. They have compared their proposed model with practical working of meteorological department and noted that their model is applicable and demonstrates reliable results for weather temperature forecasting. Likewise, Aksoy and Dahamsheh (2009) have designed three different ANN as feed forward back propagation, radial basis function and generalized regression type model with the intention of forecasting monthly precipitation in Jordan. They obtained climate data from three different meteorology stations and found out that in two stations the feed forward back propagation model outperformed and in one station other three models which are radial basis function and generalized regression type model performed reasonable as well.

Romano et al. (2009) proposed a tsunami-forecasting model constructed using ANNs. They have generated dataset via using process based computerized tsunami model which comprise the most anticipated ocean floor rupture scenarios together with the wave heights and their respective arrival times to the land. The data generated for Sunda Arc then trained and forecasts had been made via ANN. They denoted that their proposed model is capable of forecasting tsunami with low error rates and short operating time. Lam et al (2008) have designed an ANN on purpose of forecasting the residential property price in Hong Kong. Taking the micro and

macro-economic factors in to consideration namely area, age of building, district, garden, duplex, pools, number of bedrooms, distance to schools or transportation stations etc. and gross national product, number of unemployment and underemployment, climate respectively they concluded that. They defined the real housing prices as their dependent variable and concluded that the neural network performed best with smaller sample sizes.

Ghiassi et al. (2008) designed a dynamic ANN to forecast urban water demand using data gathered from actual water distribution systems. They have used historical water demand data in order to construct the model and find out significant and accurate results with the scores of 99% for week and daily predictions and 97% for hourly water demand forecast respectively. Kisi (2007) have proposed and evaluated different types of ANN models in purpose of forecasting streamflow located in North Platte River, US. He applied four different ANN models namely as Levenberg–Marquardt, conjugate gradient, backpropagation and cascade correlation by historical water data and found out that the best performance was acquired with Levenberg-Marquardt model.

Kon and Turner (2005) evaluated ANN and basic structural methods, namely, naïve and Holt-Winters, on a comparative basis, together with the purpose of forecasting tourist arrival to Singapore. The data used were gathered from the Singapore Tourism Board, which includes the topmost tourist-generating countries, namely Australia, China, India, Japan, the United Kingdom, and the USA. They have used three different learning methods for ANN as multi-layer perceptron, radial basis function, and Bayesian. They concluded by stating although basic structural methods performed well in comparison with the conventional time-series methods model created with ANN outperformed the basic structural methods in longer time horizons. Guan and Huang (2004) have designed an ANN which forecasts incidence of hepatitis A by using past data of incidence of hepatitis A from Liaoning Province between the years of 1981 to 2001 and obtained quite satisfying predictions with reasonable error rates for study area.

Campolo et al. (2003) designed an ANN model to predict the flood occurrence on River Arno, Italy, which caused damage to billions of people previously in the area, by using two years of hydrometric data, rainfall amount, and dam operations for related basins and concluded that

their proposed model generated satisfactory results. Lee and Jeng (2002) have constructed a forecasting model with back-propagation ANN in order to make tidal forecast in Taiwan. They have used historical tide level data gathered from three locations namely as Keelung, Hsinchu and Kaohsiung harbor for three tide types as diurnal, semidiurnal and mixed and have emerged the tide forecasting potential of ANN with reasonable error rates.

Jain et al. (2001) applied six different ANN models, five different regression models, and two time-series models to forecast short-term water demand at the Indian Institute of Technology, Kanpur campus, and compared the performance of these models. They utilized the data of water demand, rainfall amount, and maximum temperature of the city of Kanpur to forecast the demand in the hypothesis framework, which states that the short-term water demand process at the campus is mainly driven by the maximum air temperature and is negatively affected by rainfall occurrences, which is a more decisive variable than the rainfall amount within this time horizon. They noted that the models run with ANN continuously outperformed in comparison with regression and time series model with 2.41% average abs error and found out that water demand in campus is more related with the rainfall occurrence rather than the amount of rainfall.

Goh (1998) have developed three different models as Box-Jenkins, ANN and multiple loglinear regression in basis of comparison. The author clustered the indicators theoretically and statistically as GDP, productivity, interest rate, population, property price index, unemployment, consumer price index, building tender price index, building material price index, housing loan, and so on, and concluded that ANN outperformed the other two models where the multiple log-linear regression was least accurate.

It would not be wrong to say from finance to meteorology, from the use of natural resources such as water to tourism, from the health sector, whose importance has once again emerged in difficult times to the estimation of natural disasters, from construction to the real estate sector, from various other manufacturing and service sectors, artificial neural networks have frequently been utilized both in practice and in the literature.

2.3. Applications of Artificial Neural Networks in Energy Sector

As in many areas, it is possible to say that artificial neural networks are frequently used in the energy sector, and reflections of this have also been reported in the literature. For instance, Somarathne et al. (2022) have constructed two forecasting models with ANN and ARIMA in order to predict short term electricity demand for Sri Lanka. They have utilized six years of historical hourly demand data and found out that ANN is also a suitable and effective approach for forecasting short term electricity demand. Saglam et al. (2022) have proposed and compared three different types of forecasting models such as ANN, MLR and particle swarm optimization which predicts electricity demand of Gokceada island located in Turkey. They used number of tourists, car and import export along with the electricity data and found out that ANN have produced higher confidence interval in comparison with models utilized in the study.

Demircioglu and Esiyok (2021) built an ANN model that forecasted the gross energy consumption of Turkey. They defined their indicators as gross domestic product, import, sales prices, temperature and population and forecasted the future energy consumption of Turkey until 2025 0.99912 R^2 value. Manobanda et al. (2022) have presented ANN which forecasts the energy demand in an oil producer firm located in Ecuador. They have utilized historical fluid production and energy data. Comparing their results also with simple linear regression, the authors suggested that ANN could be used in such estimation studies.

Puri and Kumar (2021) designed an ANN to determine the potential of wind energy in the Himalayan region. They used historical temperature, wind speed, and air density data for training in MATLAB software. The results showed significant accuracy and they also stated that forecasting is an important tool since it creates a more reliable future prospect and the model proposed have the potential of forecasting the wind energy capacity. Hamzacebi et al. (2019) have proposed an ANN forecasting model by using historical monthly electricity consumption amounts gathered from TEIAS between the years 2002 and 2014 and taking the seasonal effects in to consideration. They proposed four different models for four seasons together with a superior ANN and also built a SARIMA model for comparison purpose and found out that ANN model outperforms.

In another example, Singh et al. (2017) designed an ANN that forecasts the short-term loads separately for weekdays and weekends by considering the humidity, temperature, and electricity load of the region. They have utilized from the data which is gathered from 2004 to 2008 and achieved accurate results for both weekdays and weekends with the MAPE rates of 1.38% and 1.39% respectively. Liu et al. (2017) on same topic, have constructed three different forecasting model as multi-variable linear regression, support vector regression and gated recurrent unit in order to predict the primary energy consumption of China by using the historical energy consumption data between the years 2015 and 2021. They also define their influencing factors as the gross domestic product, population and the number of imports and exports. The authors who compared different forecasting models have concluded that the best performance has been measured with gated recurrent unit type ANN with respect to MAPE AND RMSE rates.

Wang et al. (2016) constructed a dynamic choice artificial neural network in which the trained data were optimized by the index of the bad samples matrix to forecast the electricity price. They used the historical electricity price data of Queensland, Australia, for 2010. Defining this artificial neural network, which they created with the optimization of data approach, as a hybrid model, the authors stated that they obtained more accurate results compared to benchmark models. Chae et al. (2016) have designed an ANN with Bayesian regularization algorithm in order to forecast the electricity consumption of commercial building by using different factors such as humidity, wind speed and direction, sky condition, temperature and precipitation and type. They have concluded that the forecasting model they built is capable of predicting the short-term electricity consumption in commercial buildings with accurate outputs.

Almonacid et al. (2014) constructed a DANN model that forecasted the short-term power output of photovoltaic generator. They underline the importance of renewable energy, such as solar energy, and denote its drawbacks, namely, unpredictability and uncontrollability. In this manner, the authors who conducted their study within this framework used atmospheric parameters such as solar global irradiance and temperature as factors influencing the short-term load and obtained accurate results. Rodrigues et al. (2014) have designed an ANN which associated with feed forward and Levenberg-Marquardt algorithms in purpose of hourly energy

consumption and loads of household in Lisbon, Portugal. They have used numbers of occupants, apartment areas, electrical appliance consumption and Boolean inputs as meters per hour system as input variables. They have also utilized from historical consumption records between 2000 and 2001 which covers both weekends and weekdays. They have found out that their model is capable of forecasting the hourly energy consumption and loads of households accurately.

Voronin and Partanen (2014) proposed a hybrid model that combines ARIMA, ANN, and wavelet transforms to forecast the electricity price and demand simultaneously. They used historical hourly electricity consumption data from the Finnish market and obtained accurate results compared to using separate models. Kou et al. (2013) have approached photovoltaic energy output prediction a little differently than conventional models and integrate the aerosol index as well to the factors affecting energy output. On purpose of train the data they have used historical meteorological data and average output power data following that by using the weather data of objective day they have forecasted the energy output. They stated that their novel model which includes aerosol index have outperform the conventional model which is not taking the aerosol index in to account with incontrovertible MAPE results.

Buhari et al. (2012) proposed an ANN that forecasts the short-term load of a substation located in Nigeria. They utilized the daily load data of this station collected by a local utility company, which included data from 2005. On purpose of training the data Levenberg-Marquardt technique has been used together with the back-propagation algorithm. They obtained satisfactory forecasts and noted that their proposed model is capable of forecasting short-term loads which could also be beneficial for taking decisions in operational level. Accordingly, Kandananond (2011) have proposed forecasting models which approaches the phenomenon by different techniques namely ARIMA, MLR and ANN. The model utilizes from different factors such as gross domestic product, population, stock index, revenues from exported products and historical electricity consumption data of Thailand which covers the years between 1986 and 2010. With respect to MAPE rates ANN have outperformed ARIMA and MLR models respectively in the study. However, the author concluded that there is no significant difference between these models by taking the paired test results in to the account.

Singhal and Swarup (2011) designed an ANN that forecasts electricity prices using different factors such as the historical market clearing price, system loads, and fuel prices. They defined the ANN model as a three-layer backpropagation network and concluded that their proposed model operated accurately and was strongly correlated with the clearing price and load demand. Cunkas and Altun (2010) have designed two ANN namely three-layered backpropagation and recurrent neural network in purpose of forecasting the long-term energy demand of Turkey. They have used gross national and domestic product, number of households, population, industrial production index, oil price, electricity price and consumption per capita as their influencing variables. Both models have produced accurate forecast outputs, however, recurrent neural network model performed a bit better in comparison.

In the study by Catalao et al. (2009), an ANN was designed for another renewable energy source to forecast short-term wind power in Portugal. Studies based on a similar framework, such as unpredictability and volatility, have also applied persistence and ARIMA analyses for comparison purposes. They have directly utilized from the historical wind power data and ANN model outperformed other models namely persistence and ARIMA with obvious MAPE rates for all seasons. Wang and Liang (2009) have constructed an ANN with back-propagation algorithm in order to forecast the energy demand in China. They have used population, gross domestic product, level of urbanization, energy consumption, tertiary industrial output and primary and secondary industrial values as influencing factors in their forecasting model. They generally utilized the data between 1978 and 2007, however they used last two years as testing variables. They concluded that although their model is capable of producing high accuracy forecast results, due to the China's fluctuating economy after 2004 error rates increases.

Price estimation is a concept that frequently used in academia, in this context, Pino et al. (2008) have proposed a forecasting model which utilized multilayer perceptron and ART type neural network in order to forecast day-ahead electricity price in Spain. They used historic hourly electricity price and demand data from 1998 and 2004 to forecast the next day's output. They also proposed a Box-Jenkins ARIMA model for comparison purposes and found that the model constructed with ANN outperformed the conventional ARIMA model in terms of the MAPE rate, especially on weekends and holidays.

Azadeh et al. (2008) constructed an ANN with a supervised multi-layer perceptron model to forecast the long-term energy consumption of high-energy-consuming industrial sectors, such as metal and non-metal mineral industries and chemical industries. They utilized the historical energy consumption of these types of industries between 1979 and 2003 in Iran. They noted that their model produced accurate results and underlined the advantages of ANN approach for analysis of variance. The studies related with the forecasting of electricity demand or electricity consumption are usually encountered. Abdel-Aal (2008) have proposed both ANN and abductive networks for comparison purpose by using historical load and temperature data in Seattle, USA between the years of 1985 and 1992. Forecast results indicated that the neural network model superior with respect to MAPE rates. Romera et al. (2008) have designed a hybrid forecasting model by combining ANN with Fourier series in order to forecast monthly energy demand of Spain. They have used historical monthly energy consumption data and obtained satisfactory forecasts results with respect to MAPE rate.

Sozen et al. (2006) proposed an ANN that forecasts net energy consumption in Turkey by using different input variables, such as gross generation, population, and installed capacity. After statistically testing the variables, the authors concluded that the model constructed using the ANN produced accurate outcomes and underlined the importance of their study by noting the usefulness of their model in practice for planning purposes.

Kaotanzad et al. (1998) have utilized from the software namely Artificial Neural Network Short-Term Load Forecaster (ANNSTL) which is widely used in US and Canada. In addition to humidity, temperature, and wind speed, the third generation they developed, as they called it, took holidays as a factor and achieved more consistent and accurate results. Papalexopoulos et al. (1994) designed an ANN to forecast the load for the energy management system of the Pacific Gas and Electric Company. They have considered holidays, special events, heat waves and cold snaps, phenomenon which can affect normal pattern in other words, too in their model and compared it with currently used regression-based forecasting approach. As result, their proposed model continuously outperformed the existing model and noted that their model also integrated to company's energy management system.

From the literature review, it can be stated that there are an appreciable number of academic studies on forecasting energy demand consumption, and artificial neural networks are frequently preferred in these studies and produce accurate results. In these studies, it was observed that, as influencing factors, energy consumption, gross domestic product, population, temperature, number of households and enterprises or industrial businesses, gross generation, and installed capacity were highly utilized. Although some determinants differ according to the type and nature of the studies conducted, such as wind speed in wind energy. It has been found out that the same variables are generally used as influencing factors within the framework of forecasting energy consumption or demand. In this context, it is habitual that some influencing factors to be used in this study are parallel with the factors utilized in literature.

2.4. Hypothesis

As a result of the literature review, it can be stated that various factors such as population, total installed power, energy consumption, number of households and industrial facilities, and export and import amounts positively affect the demand for energy. It has been determined by expert opinions that the power transformers, subject to this study, are directly proportional to the demand for energy and energy consumption, and the production volume and number of installed PTs are defined as other factors. So, in parallel with the literature research and expert opinion the main objective of this study is investigate the following hypothesis: 1) the demand for the power transformers is driven by production volume of PTs population of subjected country and topmost exported countries, total power installed, total power generated, number of enterprises, electricity consumption of topmost exported countries and subjected country and number of installed PTs on subjected country.

3. METHODOLOGY

In this section, more comprehensive information about artificial neural network methodology is proposed together with the historical background and development of methodology and the linkage between neurology science. Furthermore, the architecture of artificial neural networks or basic components of the network in other words, have been introduced in conjunction with learning rules and strategies respectively. Conceptual information has been summed up with the evaluation methods of the outputs, results i.e.

3.1. Historical background of Artificial Neural Networks

On the way from primitive machines to the cutting-edge computers of our era, human curiosity and desire to explore played a major role. In today's world, computers that only perform arithmetic calculations in our recent past are asked to learn from events and make decisions by considering environmental conditions. One of the most important factors triggering this demand is indisputably the use of artificial neural networks in practice. The emergence of artificial neural networks has similarly started with the increasing interest of human beings in neurobiological studies. However, it is possible to state that valid studies about artificial neural networks in terms of engineering were carried out in the last quarter of the 1900s (Öztemel, 2020).

One of the facts that most augments the concept of artificial neural networks is the multidisciplinary studies in psychology, neuropsychology, and physics by connoisseurs such as Pavlov, Helmholtz, Mach, and others. These studies have played a significant role in human in manners of understanding of learning, comprehension and causation, whereat, considering that artificial neural networks use algorithms similar to human learning, the rise of this concept was accompanied by advances in these fields. (Anthony & Bartlett, 2011)

The foundations of artificial neural networks, which are still used in engineering, were laid with the work of scientists, namely McCulloch, Pitts, and Hebb. Researchers, who put forward the idea that cells work in association with each other, shaped the Hebbian learning rule, which still

forms the basis of most learning algorithms today, and paved the way for formulating many logical expressions with artificial neurons. (Öztemel, 2020)

Subsequently, mathematical model research based on biological neurons was carried out, and new structures and learning algorithms were developed and proposed. For example, Rosenblatt developed the first neurocomputer and framed the conventional perceptron model in the late 50s. This study paved the way for many future studies and Widrow and Hoff laid the foundations of adaptive linear element (ADALINE) and multiple adaptive linear element (MADALINE) networks, which works with the delta learning principle and still frequently used even on today's contemporary studies. The studies lost momentum in the 70s, and it is possible to state that one of the biggest reasons for this was the work called "An Introduction to Computation Geometry", in which Minsky and Papert stated that artificial neural networks were very limited and they had a lack of classifying nonlinearly divisible patterns. (Nunes Da Silva et al., 2018)

The milestone works done so far that time chronologically have been listed by Öztemel as follows:

- 1890–First publication on the human brain and its functions.
- 1911. Adopting the idea that the components of the human brain are composed of neurons with specific mechanisms.
- 1943–Introduction of the theory of computation based on artificial neurons and development of threshold logic devices.
- 1949–Developing biologically possible learning procedures to be performed by computers.
- 1956-1962- Development of ADALINE and Widrow Learning Algorithms.
- 1957-1962- Development of perceptron.
- 1965- Publication of the first machine learning book.
- 1967-1969- Development of some advanced learning algorithms (Grosberg).
- 1969–Demonstrating that a single-layer perceptron is not capable of solving problems.
- 1969–Defense Advanced Research Projects Agency (DARPA) withdrawal from artificial neural networks and support for other artificial intelligence studies. (p. 39)

As a consequence of these withdrawals and technical difficulties, such as increasing the calculation duration due to growing layers and neurons, studies on artificial neural networks have slowed down and were interrupted; nevertheless, a limited number of scientists continued to work, such as Kohonen, Amari, Cooper, Anderson, Fukushima, Grossberg, Hopfield, et al., and by developing technology in the computer field and increasing computational capability, have resulted in the rise of artificial neural networks on 80s and 90s (Heaton, 2015). The rapid development of computer technologies, especially after the 90s, has led to a significant increase in the use of artificial neural networks. Since then, numerous studies have been conducted and articles have been written using artificial neural networks, as seen in the literature review chapter on previous pages. In this context, the dates that can be considered as the cornerstone of this concept, from the discontinuation of the financial support to artificial neural networks to the 90s, are similarly listed by Öztemel as follows.

- 1969-1972- Development of linear associators.
- 1972- Development of correlation matrix memories
- 1974- Upgrades in early studies of backpropagation (multilayer perceptron) models.
- Development of unsupervised learning
 - 1978- Designing of ART model
 - 1982- Development of Kohonen learning rule and SOM model
- 1982- Development of Hopfield networks
- 1982- Development of multilayer perceptron
- 1984- Formulation of Boltzmann machine
- 1985- Development of multilayer perceptron with delta learning rule
- 1988- Development of RFB model
- 1988- Development of PNN model
- 1991- Development of GRNN model
- .. and since that they numerous models and development have been occurred on artificial neural networks (p.41)

3.2. Concurrence with Neurology

Artificial neural networks, one of the most widely used artificial intelligence methods, are a very simple imitation of the human nervous system. People receive information from the outside world with their nervous system, evaluate it by processing, and take action by making decisions. The nervous system plays a vital role in the human anatomy in receiving, processing, and evaluating information. Human learning and gaining experience are realized by the nervous system and the most basic structural element of this vital system is called neurons. In this context, getting to know this structural element closely is very significant in terms of understanding artificial neural networks phenomena (Eğrioğlu et al., 2019). Below figure points at biological neurons as follows;

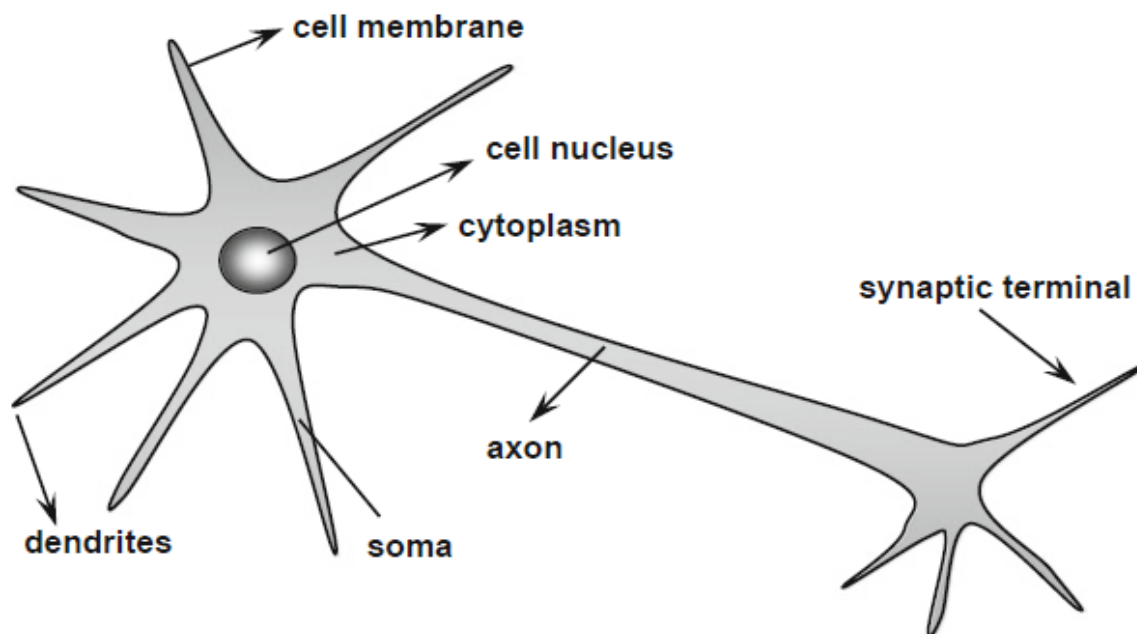


Figure 3.1 Biological Neuron

The components of biological neurons can be divided into three main parts: the cell body (or soma), dendrites, and axons. The main objective of dendrites is to receive continuous stimuli and information from the external environment via thin extensions. The information transferred from dendrites to cell body in where the information inspected and evaluated for the necessity of producing potential activation. If such a necessity arises, the electrical impulse is triggered in the body to the axon whose mission is transfer those impulses to the other connected neurons or directly to the muscular tissues called efferent neurons. Axon termination is also equipped with branches namely synaptic terminals whose objective is to relocate electric axon impulses to particular neuron to dendrites of other neurons. Although it may sound simple at first glance,

the average of 10^{11} neuron can produce average of 600 trillion synaptic connections without a physical interconnection (Nunes Da Silva et al., 2018).

Artificial neural networks are simpler than this complex structure. In these networks, the dendrites correspond to the inputs, the core of the main body, the nucleus or soma, correspond activation function, and the axon corresponds to the outputs. Considering that there are Glial cells, the number of which even exceeds 10 times of neurons it could be stated that increasing number of scientific studies in the field of neurobiology have a very high potential to make positive contributions to artificial neural networks (Eğrioğlu et al., 2019).

3.3. Neural Network Architecture

Although there are many artificial neural network models in the literature, most of them principally have similarities with each other. They consist of input layers, hidden (intermediate) layers, and output layers. Although the independent structural elements in these layers are called nodes, units, etc. in the literature, the most commonly used form of these structural elements is neurons, which are artificial neural network models inspired by human anatomy (Heaton, 2015). The following figure shows an example of a neural network with four layers: input, hidden layer 1, hidden layer 2, and output layer.

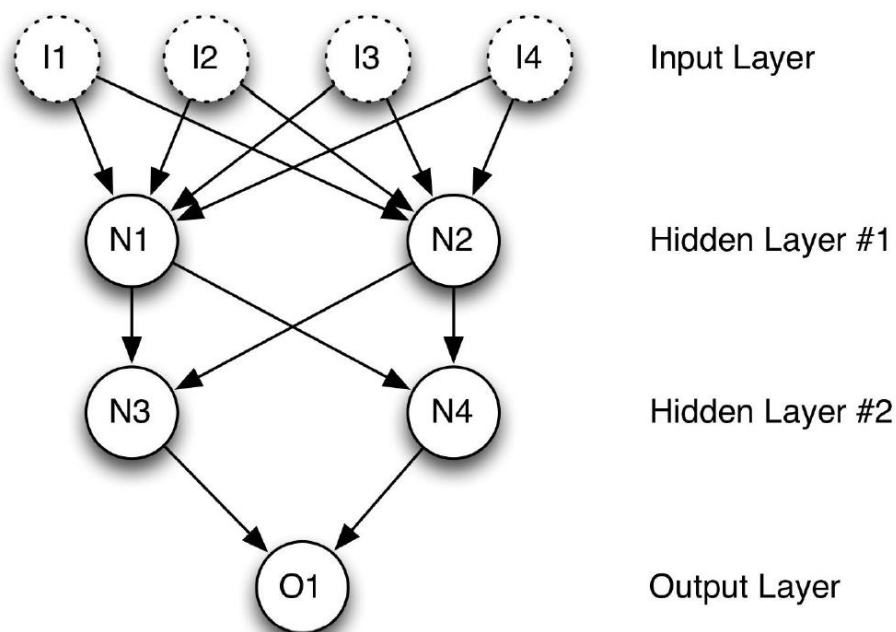


Figure 3.2 Basic Neural Network Diagram (Heaton, 2015)

Moreover, as can be seen, there are independent neurons belonging to that layer in each layer, and all of these neurons are connected to all neurons in the next layer. Connections between neurons are provided with different weight values (Hagan et al., 2014), which are called synaptic weights. (Heaton, 2015)

Neurons in the hidden and output layers of this network also have activation (transfer) functions. Although there are many variations of these activation functions, the sigmoid function is the most frequently used in the literature. This function collects the stimuli from the previous layer and compresses the output to be produced by itself to the range of 0 and 1 and transfers it to other process units of next layers (Hagan et al., 2014) with following formula:

$$\phi(x) = \frac{1}{1+e^{-x}} \quad (3.1)$$

As can be seen in the graph below, the value that is not in the range of 0 to 1 is forced to be compressed into this range by this function.

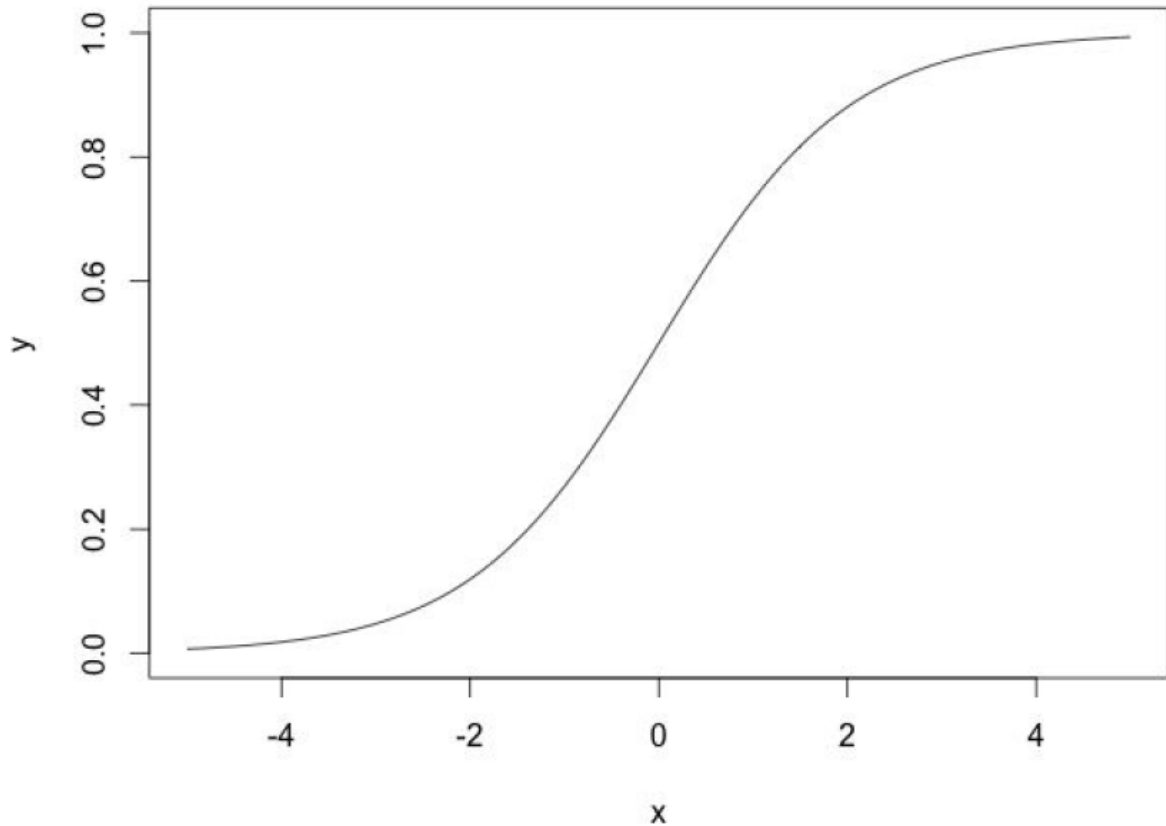


Figure 3.3 Sigmoid Function (Heaton, 2015)

One of the most important reasons why sigmoid activation functions have been utilized frequently in both literature and practice is that this function can be seamlessly derived during the training phase of the neural network. However, it is possible to use other activation functions, and the most frequently utilized activation functions in artificial neural networks are summarized in the table on the next page as:

Table 3-1 Commonly Used Activation Functions (Hagan et al., 2014)

Name	Input/Output Relation	Icon
Hard Limit	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$	
Symmetrical Hard Limit	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$	
Linear	$a = n$	
Saturating Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$	
Symmetric Saturating Linear	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = 1 \quad n > 1$	
Log-Sigmoid	$a = \frac{1}{1 + e^{-n}}$	
Hyperbolic Tangent Sigmoid	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	
Positive Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$	
Competitive	$a = 1 \quad \text{neuron with max } n$ $a = 0 \quad \text{all other neurons}$	

Neurons can also be classified according to their position in the neural network and whether they process the information. The neurons in the input, hidden (intermediate), and output layers are called input, hidden, and output neurons, respectively, as the name suggests. However, in some models, it is possible to see bias neurons in the input layer and hidden layers. Although these neurons are not located in the output layer, they do not establish a connection between themselves. The connection established is only with neurons in the next layer, which are the process elements (Heaton, 2015). A four-layer network containing three bias neurons could be illustrated as follows:

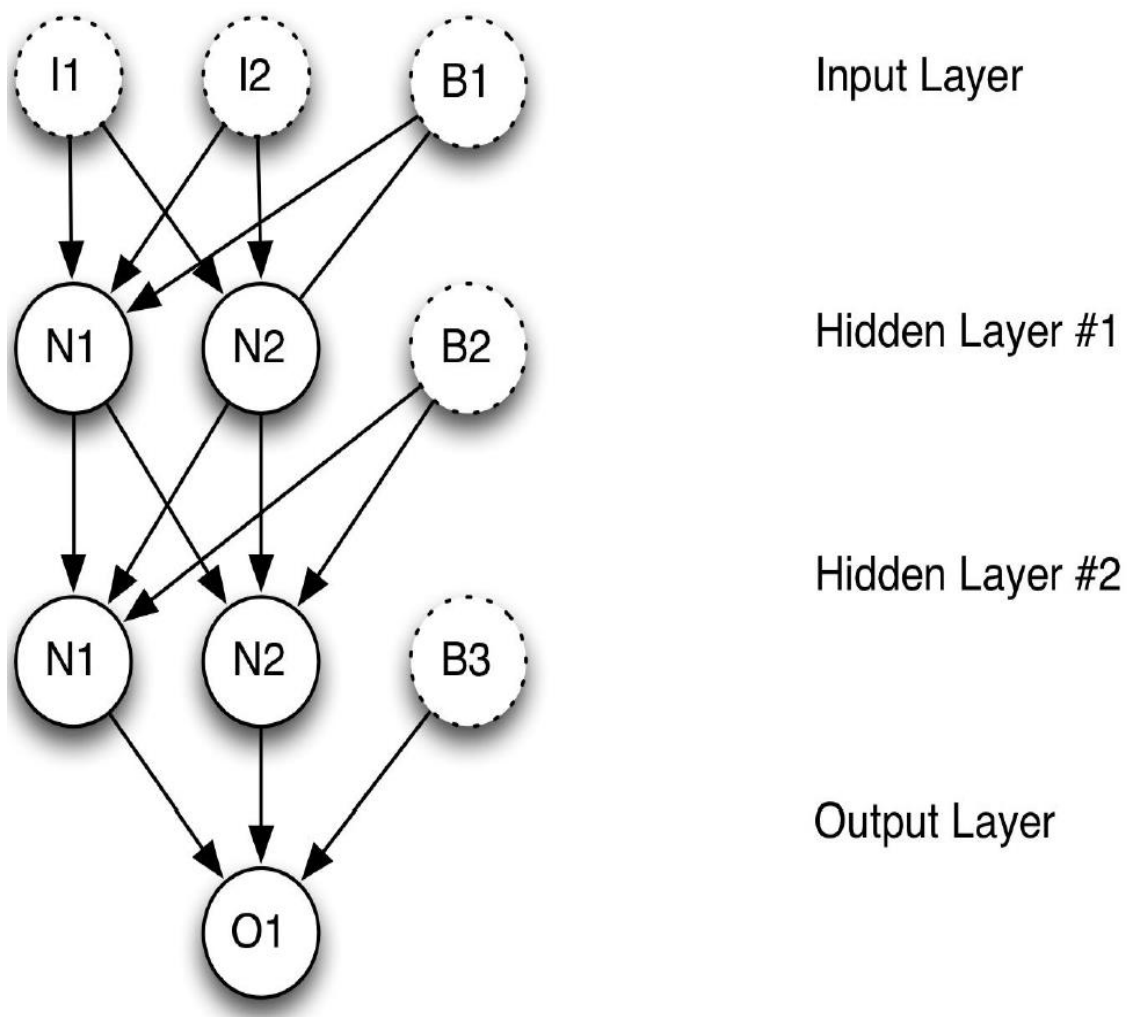


Figure 3.4 Four-Layer Network Containing Three Bias Neurons (Heaton, 2015)

As can be seen above, there are single bias neurons, only one in each layer, except for the output layer.

Bias neurons play an important role in producing more consistent results in the artificial neural network. It is possible to say that the shape of the activation function is determined by the weights and bias in the neuron. By expanding and reformulating Eq. 3.1. With respect to the bias neurons and weights, the following equation can be obtained:

$$\phi(x, w, b) = \frac{1}{1+e^{-(wx+b)}} \quad (3.2.)$$

where w and b refer to the synaptic weights of the corresponding neuron and the bias, respectively. The following figure shows the shape of several sigmoid functions without utilizing the bias neurons in the neural network.

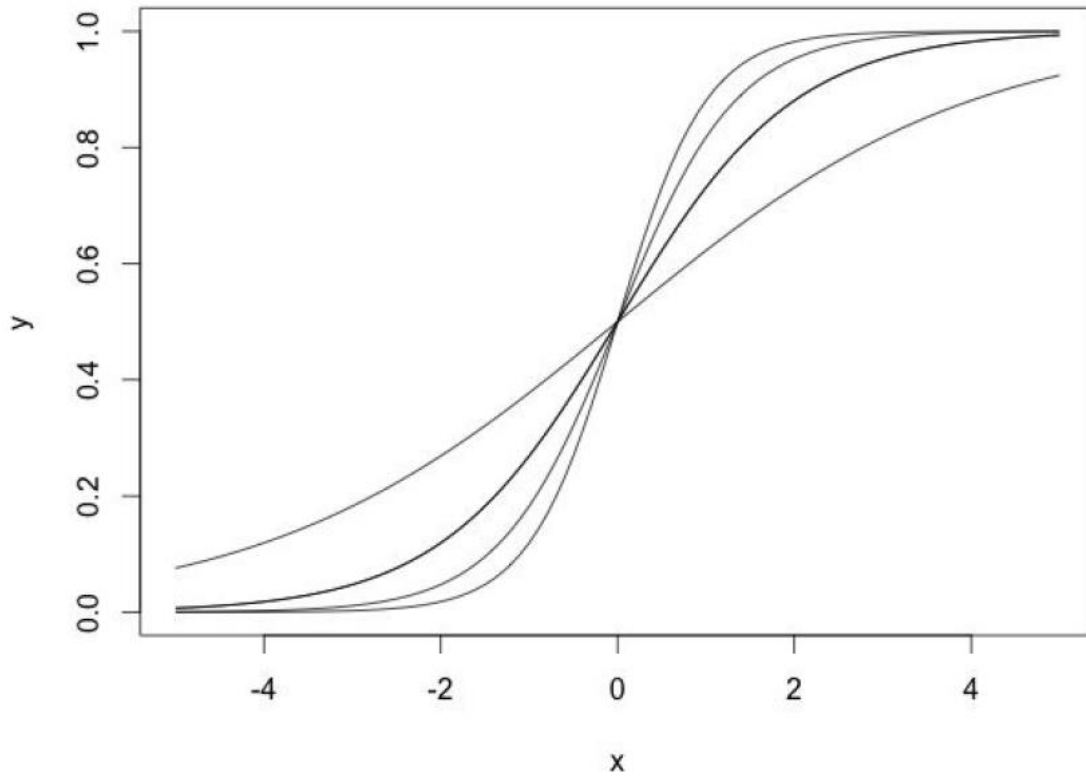


Figure 3.5 Sigmoid Curves without Bias Neuron (Heaton, 2015)

where the parameters are as follows.

$$\phi(x, 0.5, 0)$$

$$\phi(x, 1.0, 0)$$

$$\phi(x, 1.5, 0)$$

$$\phi(x, 2.0, 0)$$

The following figure shows the shape of several sigmoid functions utilizing the bias neurons in the neural network as follows:

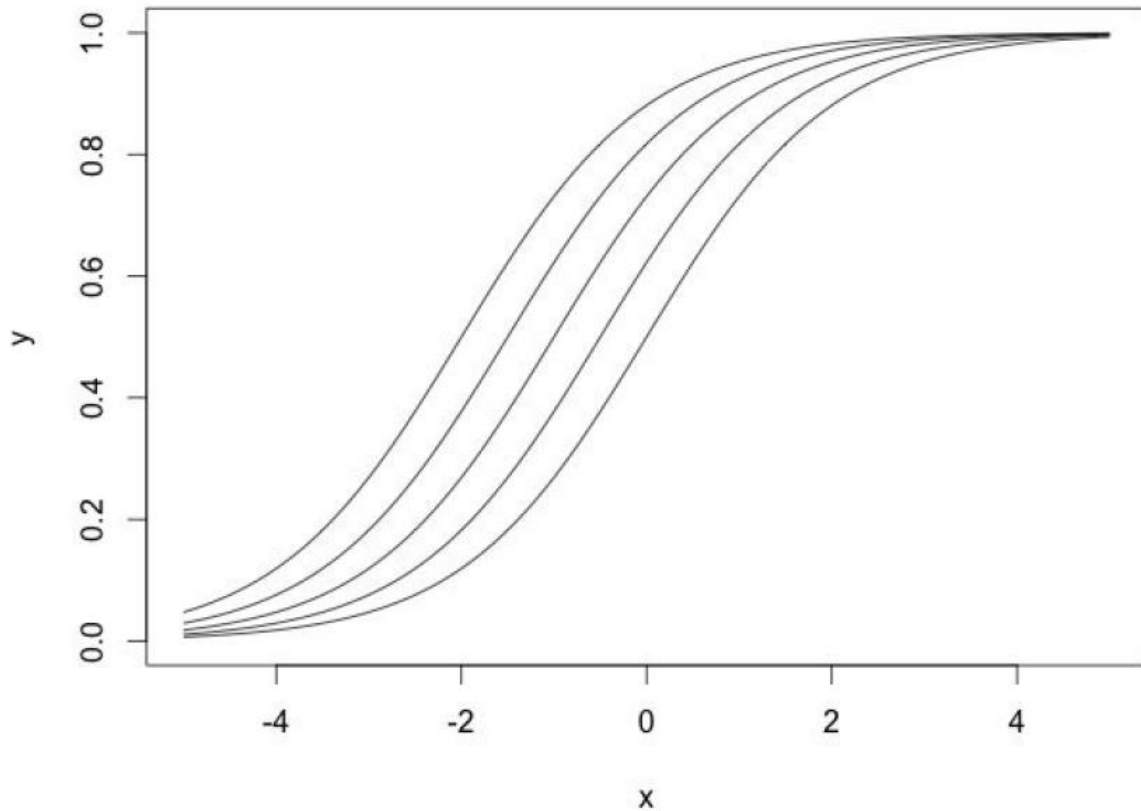


Figure 3.6 Sigmoid Curves with Bias Neuron (Heaton, 2015)

where the parameters are as follows.

$$\phi(x, 1.0, 1.0)$$

$$\phi(x, 1.0, 0.5)$$

$$\phi(x, 1.0, 1.5)$$

$$\phi(x, 1.0, 2.0)$$

As seen above, in an activation function of a neuron where all weights are the same and bias differs from 0.5 to 2.0, curves merge either at the top right or bottom left, which prevents inaccurate outputs, as shown in Figure 3.5, in which the outputs are always the same regardless of the weight when $x=0$.

To explain a simple neural network with a mathematical example, below is an artificial neural network modeling the conjunction "and," in which 1 and 0 refer to true and false, respectively, and the network produces 1 if both propositions of logical statements are equal to 1, as in (1,1).

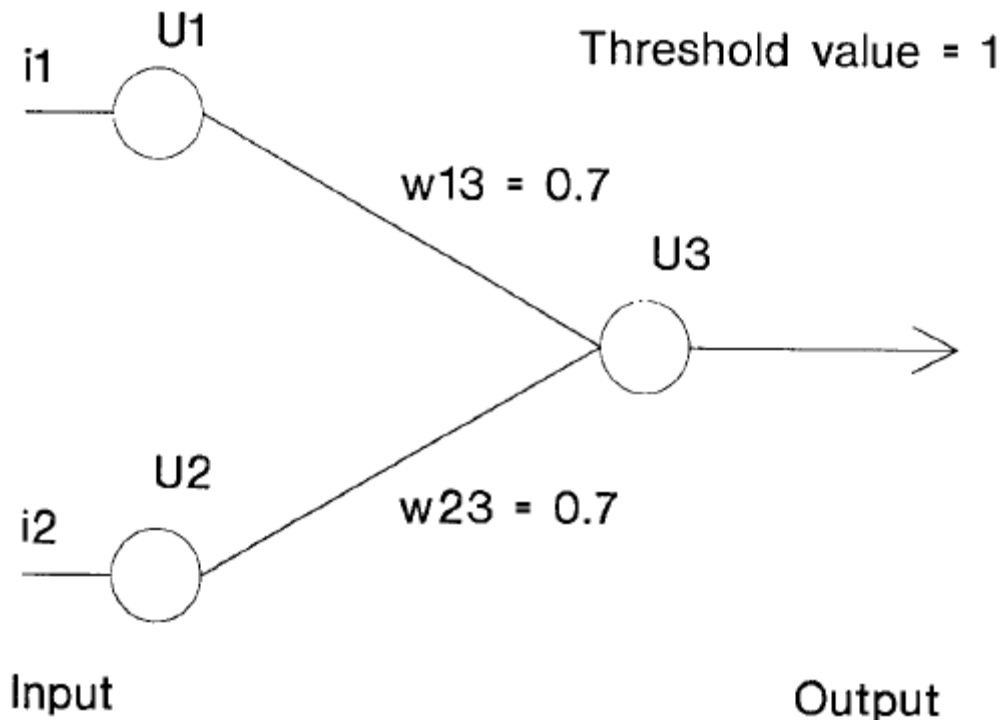


Figure 3.7 The conjunction AND (Braspenning et al., 1995)

The figure above embodies three processing units, U1, U2, and U3, the so-called input and output neurons, respectively. Weights placed on the interconnections of each part of this process units. When a stimulus transmits to the network through i1 and i2 units the signals distributed to U3 by U1 and U2 units with respect to synaptic weights of 0.7 for both interconnection and following mathematical expression could be expressed:

$$w1 * i1 + w2 * i2 = 0.7 * 1 + 0.7 * 1 = 1.4 \quad (3.3.)$$

Because the threshold value is equal to 1, the logical expression produces the value 1, which means true via the output unit. If it's assumed that $i1 = 0$ and $i2 = 1$ the network produces 0 which means false since the calculation of below equation is smaller that threshold value 1.

$$w1 * i1 + w2 * i2 = 0.7 * 1 + 0.7 * 0 = 0.7 \quad (3.4.)$$

As a result, a very simple artificial neural network with a single perceptron performs according to the principles mentioned above.

3.4. Multilayer Perceptron

A multilayer Perceptron is a neural network in which at least one hidden (intermediate) layer is placed between the input and output layers. MLP, one of the most widely used neural networks in literature and theory, is a versatile application. Some of the application areas of MLP could be listed as follows;

- Curve fitting
- Pattern recognition.
- Process identification and control.
- Forecasting.
- System optimization.

For a general classification, MLP is a supervised feedforward approach and one of its most important features is its ability to solve nonlinear problems that hinder the progress of artificial neural networks in the past. Below figure illustrates a MLP with two hidden layers with n hidden neurons together with output and input layer (Nunes Da Silva et al., 2018; Öztemel, 2020)

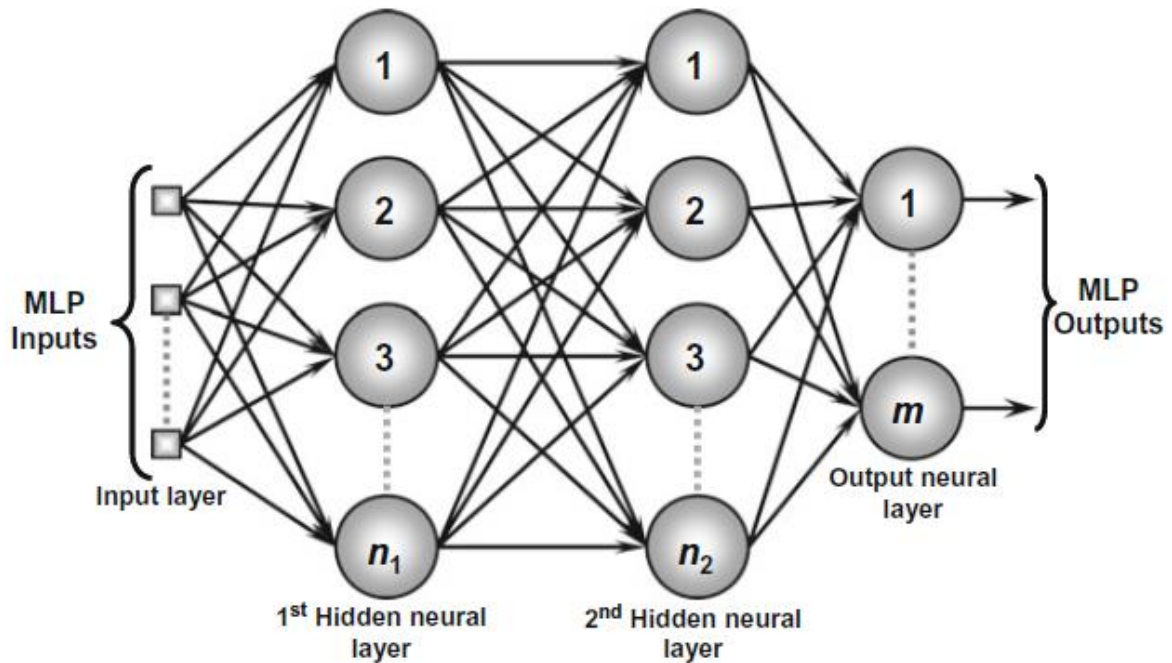


Figure 3.8 Multilayer Perceptron Network Diagram (Nunes Da Silva et al., 2018)

The main objective of the multilayer perceptron is to calculate the weights that produce the outputs closest to the target values. While there is no connection between neurons in the same hidden layer, connections between layers are forward-only. Thus, calculation of outputs is called feedforward in multilayer perceptron. In feedforwarding the layers on artificial neural network operates by turns (Eğrioğlu et al., 2019). This situation is the same between hidden layers containing neurons, and with different perspective an ANN consisting of two hidden layers, the inputs of the second hidden layer consist of neurons of the first hidden layer (Nunes Da Silva et al., 2018).

In the MLP model, the main structural elements can be listed as the input, hidden (intermediate), and output layers. The input layer receives stimuli from the outside world and directs them to the intermediate layers. Although there is no information processing in this layer, every incoming information is transferred to the next layer as it is. The signals reaching the input layer are transferred to all the intermediate layer elements, namely neurons. In other words, it is possible to say that every process element in the input layer is connected with all the process elements in the next layer.

In the intermediate (hidden) layers, the situation is almost the same, except that the information coming from the input layers is processed and transferred to the next layer. As noted earlier, an MLP network can have more than one hidden layer and more than one neuron belonging to these layers. Each process element in the intermediate layer is connected to all the next process elements, just like in the input layer.

In the output layer, which is the last structural element of the MLP network, the incoming information from the last intermediate layer is processed and the outputs produced by the network are transferred to the outside world in response to the data transferred from the input layer of the network. Similarly, more than one process element may exist in the output layer. Each process element is linked to all process elements in the previous layer. However, all process elements in the output layer have only one output (Öztemel, 2020).

In summary, in a multilayer perceptron, information is presented to the network from the input layer and reaches the output layer by passing through the intermediate layers. In response to the inputs presented to the network, the network's response is transmitted to the outside world and the operation is completed.

3.5. Training Processes and Learning Strategies

Undoubtedly, one of the most important factors in the prominence of artificial neural networks is their potential to learn from inferences made from samples. Thanks to these features, these networks can mathematize the relationship between the inputs and outputs and provide more consistent outcomes on operations such as; pattern recognizing, optimization, forecasting, etc. (Nunes Da Silva et al., 2018). The training process in artificial neural networks is divided into two in general terms as supervised and unsupervised learning. With another classification, training methods are similarly may divided into two as online and offline learning.

In supervised training, the trained system is supervised by an external phenomenon, as suggested by the name. The supervisor provides samples of the event to be learned to the system as input and output sets. In other words, for each sample, both the inputs and the outputs that need to be generated with respect to those inputs are introduced to the system. The main objective of the neural network is mapping the inputs to the outputs specified by the supervisor which makes it possible for network to be trained by considering the relationship between the inputs and outputs (Öztemel, 2020). This learning strategy was pioneered by Donald Hebb, (1949) inspired by neurophysiological observations. Multilayer perceptron used in this study is also utilizes from this type of training algorithm.

Unlike supervised training, unsupervised training does not require any information regarding the targeted outputs. Therefore, the neural network needs to determine the relationship between the components of the network by creating subsets or clusters with respect to similarities (Nunes Da Silva et al., 2018). In this type learning algorithms only the input variables introduced to the network by its' designer and additionally labels representing what the outputs mean after the system has learned should also been introduced the system by network designer (Öztemel, 2020).

The two learning strategies, supervised and unsupervised, can be defined as iterative processes. Utilizations of these learning strategies differ according to the sample and the type of the problem. In the supervised learning strategy, the outputs are introduced to the system together with the inputs and the network is expected to converge to the minimum error, while in the unsupervised learning strategy, the outputs are not introduced to the artificial neural network and it is difficult to calculate the resulting error, so the designer of the network is expected to make sense of the outputs (Heaton, 2015).

If learning strategies are to be classified from another perspective, they can be divided into online and offline learning strategies. In learning systems such as artificial neural networks, learning takes place according to certain rules, regardless of which strategy is used, and some of these rules work online and some offline. Online learning rule is a learning approach that works by real-time. Systems that learn with these rules, while working in real time, perform their functions on the one hand and continue to learn on the other hand (Öztemel, 2020). This learning strategy re-determines the threshold and weight values after each updated sample and produces new outputs. In this context, they are used in network models whose behavior frequently changes. However, although it is a more applicable method in practice, it is thought that it encounters difficulties in producing consistent results without increasing the size of the sample that the neural network utilized from (Nunes Da Silva et al., 2018). The learning rule of the ART network and the Kohonen learning strategy could be shown as examples of online learning strategies (Öztemel, 2020).

In offline learning, the calculation of the weights and threshold values can be realized after the entire training set is introduced to the neural network because the errors calculated through the outputs are used in the recalculation, adjustment, and approximation of these values. In this context, offline learning strategies should make at least one epoch so that the mentioned values can be calculated (Nunes Da Silva et al., 2018). Artificial neural networks using this type of learning strategies are no longer learns after they have been trained and put into practice. When new information that is thought to be learned by the system emerges, the system is taken out of use till whole the process ended and retrained as offline. The delta learning rule, which is

frequently used in the literature, uses this type of learning strategy and could be given as an example for offline learning strategies (Öztemel, 2020).

In summary, in systems with learning capabilities such as artificial neural networks, the types of learning strategy may vary owing to the available data set, the type of application, and the inherit of the problem. In this sense, the aforementioned facts should be carefully considered by the person who designed the network to obtain more accurate outcomes. While online learning strategies can be used in flexible systems that are suitable for rapid changes, unsupervised learning strategies can be used as a solution approach to problems such as clustering. However, it is worth noting that different learning strategies may bear even reasonable outcomes for improper matchings due to the high potential of artificial neural networks.

3.5.1. Training Processes of Multilayer Perceptron

The training phase of the MLP is similar to that of the other networks. The aim of this procedure is to produce the expected output (target values) using input samples introduced into the neural network. MLP networks, as mentioned, utilize from a supervised offline method called back propagation, also known as generalized delta learning rule, as the learning algorithm. This rule is basically enforced by the successful realization of two stages as feedforward and back propagation.

In the first stage of this methodology, $x_1, x_2, \dots, x_n$ signals are inserted into the network, which have been introduced to the artificial neural network by the supervisor, as noted earlier, by input layers, and they reach the output layer by passing through all the available intermediate (hidden) layers in the network, producing a potential outcome. Subsequently, on second stage network back propagates the output for the purpose of minimization of total error (Nunes Da Silva et al., 2018). Following figure illustrates an information flow and feedforward stage and backpropagation respectively in a multilayer perceptron network in which there is n number of signals and two hidden layers;

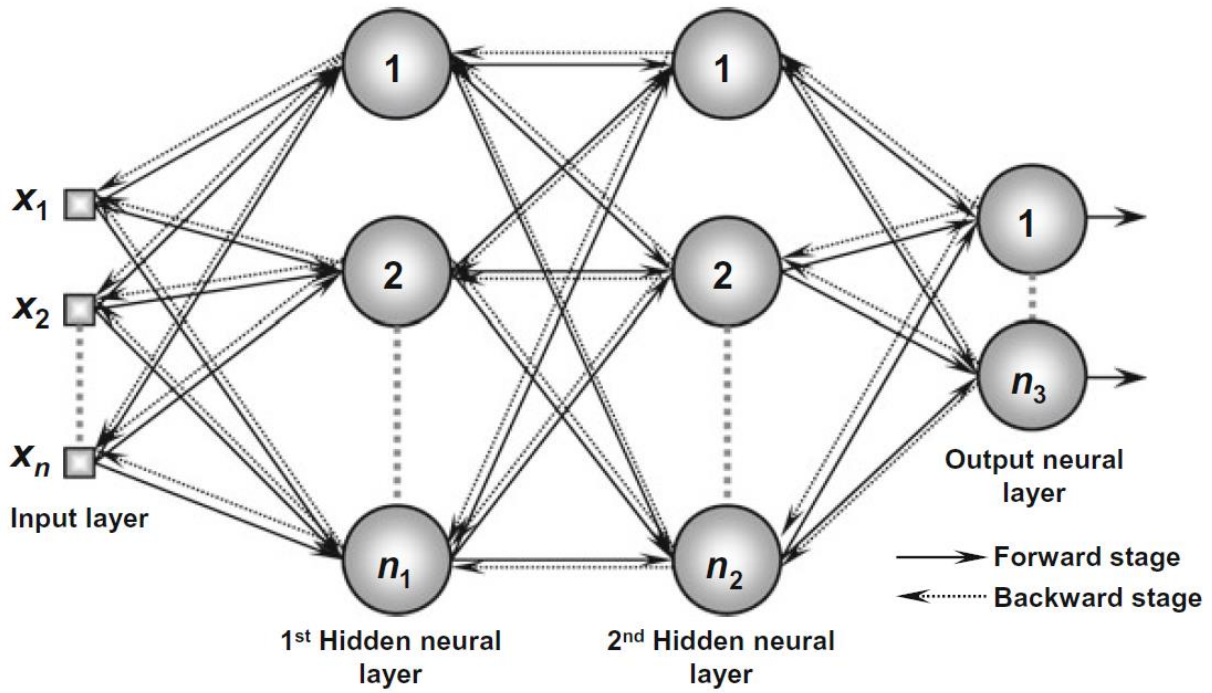


Figure 3.9 Information Flow in an MLP Network (Nunes Da Silva et al., 2018)

In the first stage of learning, that is, in the input layers, no information processing is performed. As mentioned before, the inputs inserted into the system by the network designer are transferred to each process element of the hidden layers as it is, without any changes. The output of the k^{th} process element in the input layer can be mathematically expressed as:

$$C_k^i = X_k \quad (3.5.)$$

Each process element in the intermediate layer receives synaptic connection weights from all the process elements in the input layer. The net input to the process elements in the intermediate layers, namely, neurons, can be expressed as follows:

$$NET_j^a = \sum_{k=1}^n A_{kj} C_k^i \quad (3.6.)$$

where A_{kj} is the synaptic weight of the connections between the k^{th} input layer element and the j^{th} hidden layer element.

As previously mentioned, information processing occurs in the neurons of the hidden layers. In this context, the output of the j^{th} hidden layer element is calculated by processing the aforementioned equation using the activation function. As disclosed earlier, activation functions vary either as hyperbolic tangent or sigmoid functions. In this study since the sigmoid function is decided to be utilized following equation which performed in hidden neurons and produce the output of the respected hidden neuron could be given as:

$$C_k^a = \frac{1}{1+e^{-(NET_j^a + \beta_j^a)}} \quad (3.7.)$$

where β_j denotes synaptic weight of the threshold element connected to the j^{th} element in the hidden layer.

The output of the threshold unit mentioned above is constant and equal to 1. The synaptic weight value is set to determine the orientation of the sigmoid function, and the network determines this value itself during training. All process elements of the hidden layers and the outputs of the process elements of the output layer are determined by calculating the NET input they receive and processing them through a sigmoid function. Once the values of the output layer ($C_1, C_2, \dots, C_n$) are found the forward propagation, first stage of learning algorithm, is completed (Öztemel, 2020).

Based on the inputs presented to the network, an error value is obtained by comparing the outputs produced by the network with the expected outputs by calculating the difference between these two values. As has been emphasized many times, the purpose of these networks is to minimize the error. The error calculated in the backpropagation algorithm is distributed to the synaptic weight values of the network, and it is aimed at reducing the error in the next iteration. This stage could be formulated as follows:

$$E_m = B_m - C_m \quad (3.8.)$$

where E_m and B_m are the errors calculated for the m^{th} process in the output layer and the expected value for the m^{th} process respectively. However, this is the error calculated for only one process element, all errors must be summed to find the total error for the output layer. In case of some error values may be negative and some positive, and to prevent miscalculation in the total error, the squares of the values are calculated and the square root of the result is taken.

$$TE = \frac{1}{2} \sum_m E_m^2 \quad (3.9.)$$

In order to minimize the total error, this value should be distributed to the process elements that cause it to occur, which implies changing the synaptic weights of the process elements. This change can be done in more than one way. These are modifying the synaptic weights between the hidden layers and the output layer, the modification between the hidden layers, and modifying the weights between the hidden layers and the input layer (Öztemel, 2020).

The first of the modification mentioned above is the recalculation of the synaptic weights between the hidden and output layers. This process could be formulated as follows:

$$\Delta A_{jm}^a(t) = \lambda \delta_m C_j^a + \alpha \Delta A_{jm}^a(t-1) \quad (3.10.)$$

where ΔA_{jm}^a is the change in the synaptic weight of the link connecting the j^{th} process element in the hidden layer to the m^{th} process element in the output layer, λ is the learning rate, α is the momentum coefficient, δ_m is the error of the m^{th} output element, and t is the iteration index. The expression δ_m in the formula could be expanded as follows.

$$\delta_m = f'(NET) \cdot E_m \quad (3.11.)$$

where $f'(NET)$ is the differentiation of the activation function NET. Assuming that the sigmoid function is utilized in the network, which is preferred in this study, the following statement emerges.

$$\delta_m = C_m(1 - C_m) \cdot E_m \quad (3.12.)$$

After calculating the change amounts, the new values of the synaptic weights in the t^{th} iteration can be expressed as:

$$A_{jm}^a(t) = A_{jm}^a(t-1) + \Delta A_{jm}^a(t) \quad (3.13.)$$

Similarly, the synaptic weights of the thresholds must be modified, and this process can be formulated as follows in where β^C refers to the threshold of the process elements in the output layer:

$$\Delta \beta_m^C(t) = \lambda \delta_m + \alpha \Delta \beta_m^C(t-1) \quad (3.14.)$$

Thus, the new threshold value in the t^{th} iteration is

$$\beta_m^C(t) = \beta_m^C(t-1) + \Delta \beta_m^C(t) \quad (3.15.)$$

Another stage in the distribution of synaptic weights is the modification of the weights between the hidden (intermediate) layers, if any, or between the hidden and input layers. As mentioned before, the network may contain more than one hidden layer, according to the preference of the network designer, and each process element of these hidden layers is connected to each other

by synaptic weights. Similarly, each connection from the input layer to the hidden layer has weights, and these weights need to be modified to minimize the error. The modification between the input layer and the hidden layer can be expressed mathematically as follows:

$$\Delta A_{kj}^i(t) = \lambda \delta_j^a C_k^i + \alpha \Delta A_{kj}^i(t-1) \quad (3.16.)$$

where ΔA^i refers to the change between the hidden and input layers.

The error indicator δ^a in the above formula is calculated as follows:

$$\delta_j^a = f'(NET) \sum_m \delta_m A_{jm}^a \quad (3.17.)$$

Similarly, considering that the activation function is a sigmoid function:

$$\delta_j^a = C_j^a(1 - C_j^a) \sum_m \delta_m A_{jm}^a \quad (3.18.)$$

After calculating the error, it is possible to calculate the amount of change using the above formula, and the new values of the synaptic weights can be calculated as follows:

$$A_{kj}^i(t) = A_{kj}^i(t-1) + \Delta A_{kj}^i(t) \quad (3.19.)$$

As in the first step of the distribution of synaptic weights, the new weights of the threshold values can be calculated as follows:

$$\Delta \beta_j^a(t) = \lambda \delta_j^a + \alpha \Delta \beta_j^a(t-1) \quad (3.20.)$$

Where the β^a is threshold values of hidden layers

Thus, the new value of the threshold values in the t^{th} iteration is

$$\beta_j^a(t) = \beta_j^a(t - 1) + \Delta\beta_j^a(t) \quad (3.21.)$$

Thus, all synaptic weights in the neural network can be changed.

The iteration is completed by performing calculations based on both feedforward and backpropagation processes, and the next iteration is started by giving a second sample to the system. The same operations are repeated until learning is complete (Öztemel, 2020).

For a network to stop learning, a certain criterion must be specified. This criterion is generally determined by the user of the neural network. Although every network has an optimum solution, it is often impossible to reach this solution using conventional methods. network attempts to obtain this solution during training, together with producing local solutions.

Failure to ensure that the result produced is optimal causes the network designer to specify a tolerance. In the graph below, while the line e expresses the tolerance, W1 and W9 weights are feasible and optimum results, respectively. If the optimization is performed, the user may complete the learning as soon as acquiring the satisfactory result.

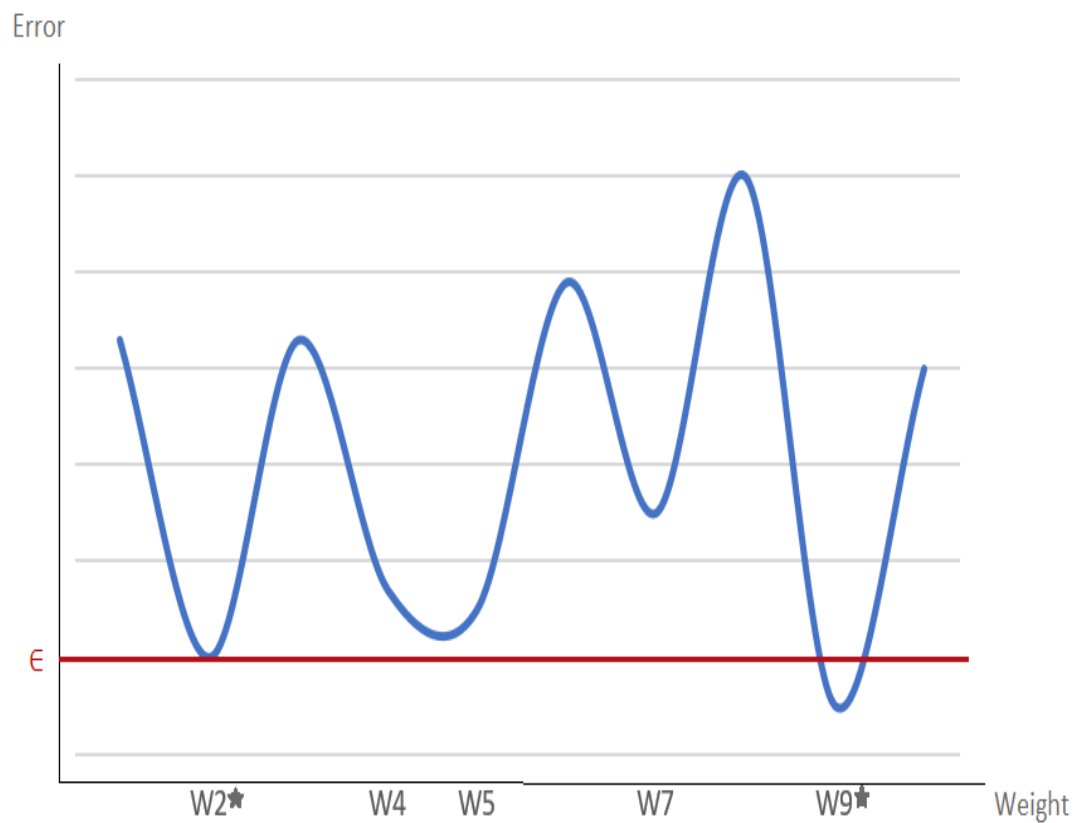


Figure 3.10 Local and Optimum Solutions Considering Weights

Additionally, to better understand the feedforward backpropagation (generalized delta rule) learning algorithm and to summarize, the heading following flow chart has been proposed on the next page as follows:

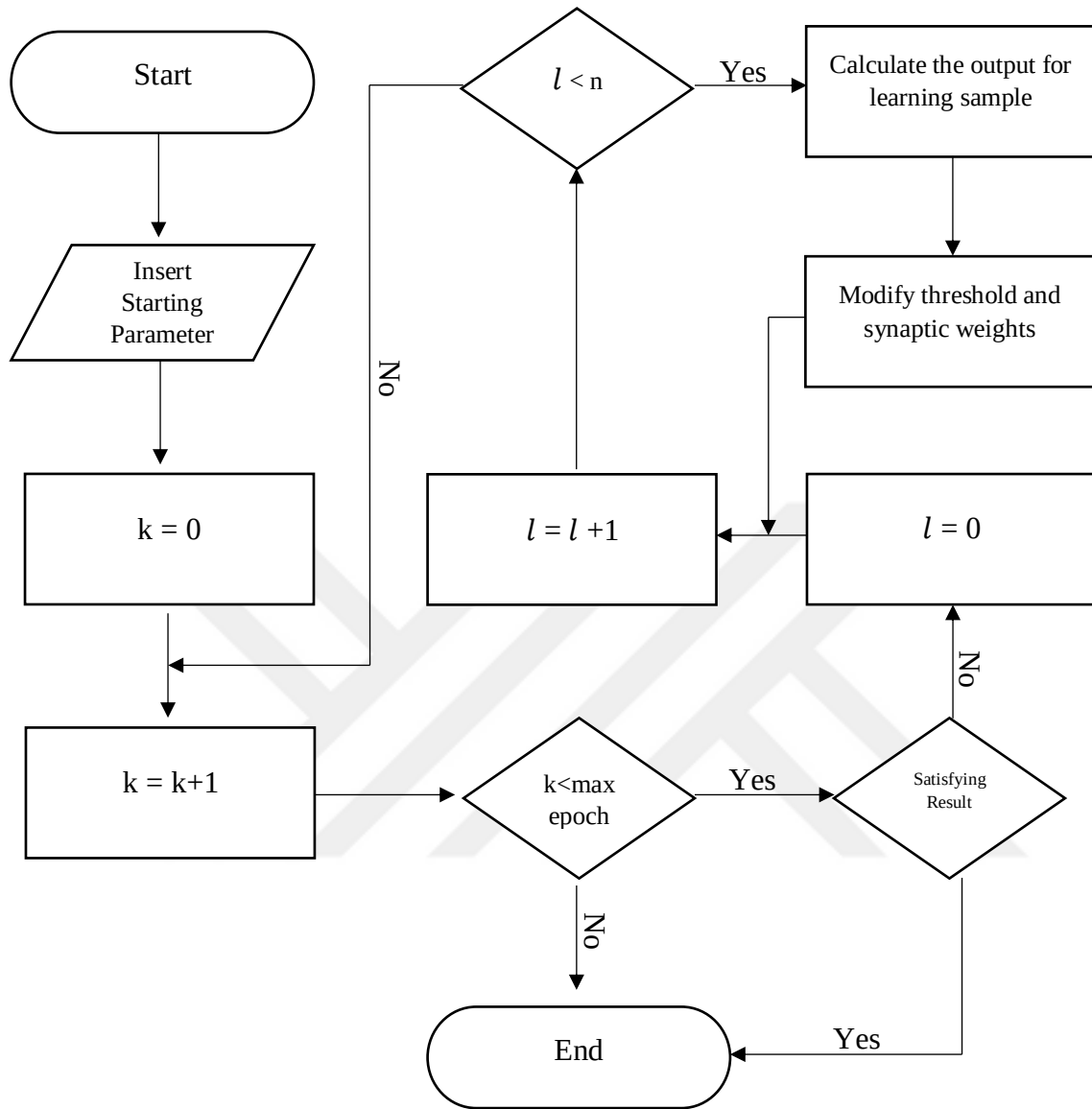


Figure 3.11 Flow Chart of Feedforward-back propagation Algorithm (Eğrioğlu et al., 2019)

Where the;

k : epoch counter

maxepoch : maximum number of epochs

l : learning sample counter

3.6. Evaluation Principle of Outputs

One of the most effective methods used to measure the performance of forecasting is statistical error metrics. Although many performance measures are used in the literature, one of the most frequently used is mean absolute percentage error (MAPE). One of the most important features that takes MAPE one step further is that anyone other than the network designer, who has no pre-knowledge about characteristics of data set, can rationally interpret this result at the first glance (Makridakis, 1993). In this sense, even if other error metrics are also used in this study, MAPE was utilized in order to measure the performance of the neural networks created while applying the artificial neural network methodology. Based on this the mentioned error metric can be formulated as follows.

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (3.22.)$$

Where the variables of

A_t : Actual Value

F_t : Forecasted Value

n : Number of points

The above equation can be modified according to the desired output format by replacing 1 with 100, which produces results in percentages. The MAPE statistic eliminates disadvantages that may arise when comparing models with different unit values. It is possible to note that since it expresses estimation errors as percentages, when the formula is modified as stated above, it is considered superior to other error metrics because it has meaning on its own. On the other hand, MAPE has some drawbacks too. For example, using MAPE in small denominators can cause problems. If there is 0 in the data set, problems will arise because there can be no division by zero. In addition, it is known that the MAPE metric carries the bias of giving more penalties to negative errors than to positive errors (Makridakis, 1993). In this context, the outputs in this study have been evaluated with many performance metrics, and some of them are summarized with their formulas on the next page.

Table 3-2 Performance Metrics

Measure	Description
Error (E)	$A_t - F_t$
Absolute Error (AE)	$ A_t - F_t $
Mean Squared Error (MSE)	$\frac{1}{n} \sum_{t=1}^n E^2$
Root Mean Squared Error (RMSE)	$\sqrt{\frac{1}{n} \sum_{t=1}^n E^2}$
Sum of Squared Error (SSE)	$\sum_{t=1}^n (A_t - F_t)^2$
Sum of Squares Total (SST)	$\sum_{t=1}^n (A_t - \bar{A})^2$
Sum of Squares Regression (SSR)	$\sum_{t=1}^n (F_t - \bar{A})^2$
R^2	$1 - \frac{SSR}{SST}$
Mean Absolute Deviation (MAD)	$\frac{1}{n} \sum_{t=1}^n AE_t$
Mean Absolute Percentage Error (MAPE)	$\frac{1}{n} \sum_{t=1}^n \left \frac{A_t - F_t}{A_t} \right $
Tracking Signals (TS)	$bias_t / MAD_t$

3.7. Implementation

3.7.1. Multilayer Perceptron Application via IBM SPSS

Under the first sub-heading of the implementation stage, one of the ready-to-use toolboxes of IBM SPSS software was used to draft and understand the primary reaction of the factors and the network. As mentioned previously in the first part of the research, nine factors were defined and used to forecast the demand for power transformers. As reminder factor could be listed as production volume, population by country, total power installed, total power generated, electricity consumption, number of enterprises, electricity consumption of topmost exported countries, population of topmost exported countries and finally number of installed power transformers. Furthermore, the demand has been converted in to a mathematical expression by associating it with the historical sales amount of power transformers. To summarize, by using the historical sales data and the factors affecting it, a forecast application has been carried out via a toolbox in this section. The mentioned variables have been listed to software as follows;

Name	Type	Width	Decimals	Values	Missing	Columns	Align	Measure	Role
pro_volum	Numeric	8	0	None	None	8	Right	Scale	Input
tr_pop	Numeric	8	0	None	None	8	Right	Scale	Input
tpi	Numeric	8	0	None	None	8	Right	Scale	Input
gross_g	Numeric	8	0	None	None	8	Right	Scale	Input
net_e_c	Numeric	8	0	None	None	13	Right	Scale	Input
no_ins_pt	Numeric	8	0	None	None	8	Right	Scale	Input
no_enterprise	Numeric	8	0	None	None	18	Right	Scale	Input
net_e_c_top...	Numeric	8	0	None	None	21	Right	Scale	Input
tmc_pop	Numeric	8	0	None	None	20	Right	Scale	Input
sales	Numeric	8	0	None	None	11	Right	Scale	Target

Figure 3.12 Data View on SPSS

In the first stage, the data were transferred to the software and the inputs were defined to the system. The inputs consist of the factors listed above that the network takes into account when forecasting. Subsequently, the outputs have been introduced to the software similarly. Outputs express what is to be estimated, and the variable to be estimated in this study is historical sales data, which also denotes the demand. After this process, independent (covariates) and dependent variables manually introduced to software and rescaling of covariates had been decided to set normalized. Normalization is a pre-processing step which utilized lot in forecasting techniques and re-scales the existent data with respect to range between variables. (Patro & Sahu, 2015) Although there are many normalization methods, min-max normalization has been utilized by software in this study. The min-max normalization forms as follows;

$$A' = \left(\frac{A - \text{min value of } A}{(\text{max value of } A - \text{min value of } A)} \right) * (D - C) + C \quad (3.23)$$

Where;

A' contains Min-Max Normalized data one

If pre-defined boundary is [C, D]

The normalization process produces outcomes within the range of -1 to 0, and closes the ranges between variables. These stages of the study are shaped through the pop-up menu in the image below.

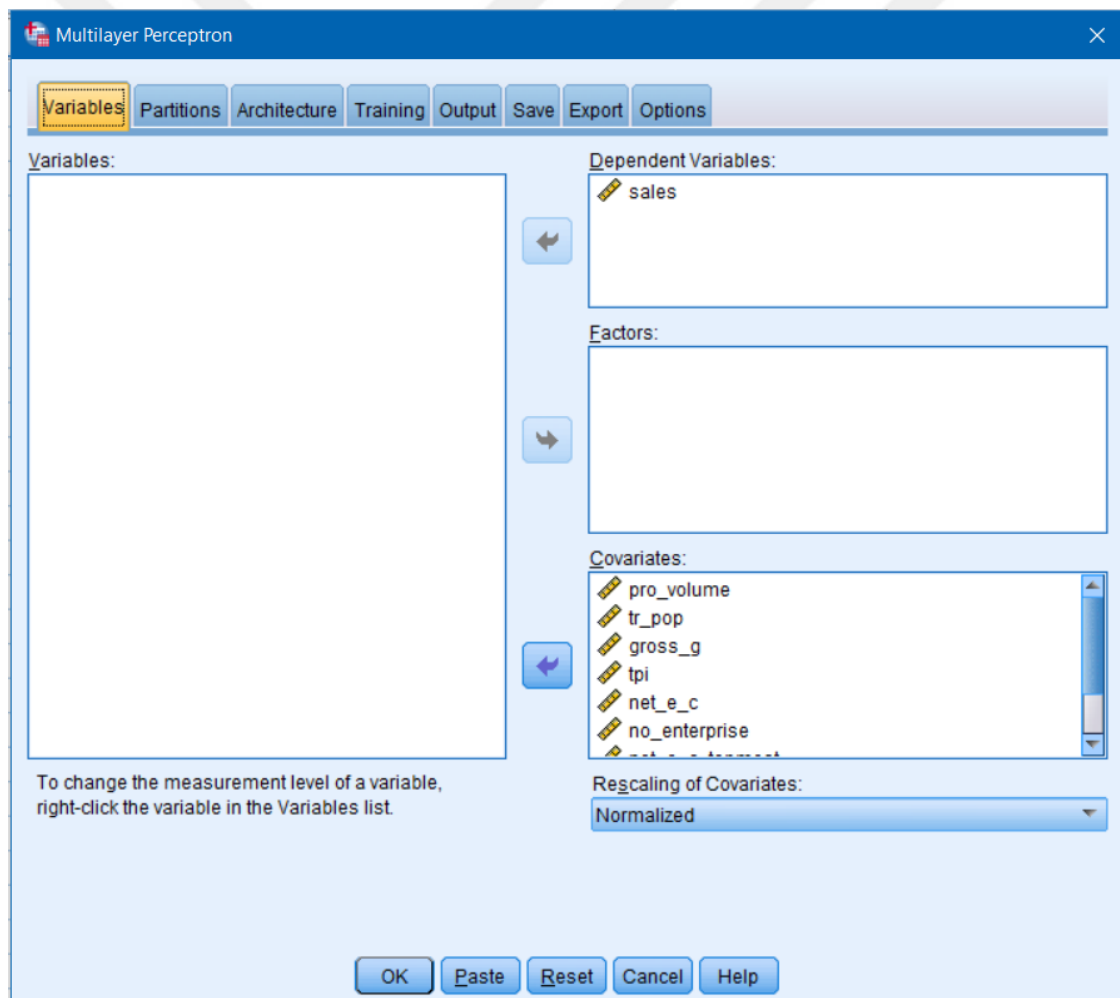


Figure 3.13 Configuration of Variables

The partition tab allows the user to decide the amount or percentage of data to be used for training, testing, and validation, respectively. Although it is a parameter that affects the performance of the network, the numbers are set as default as follows:

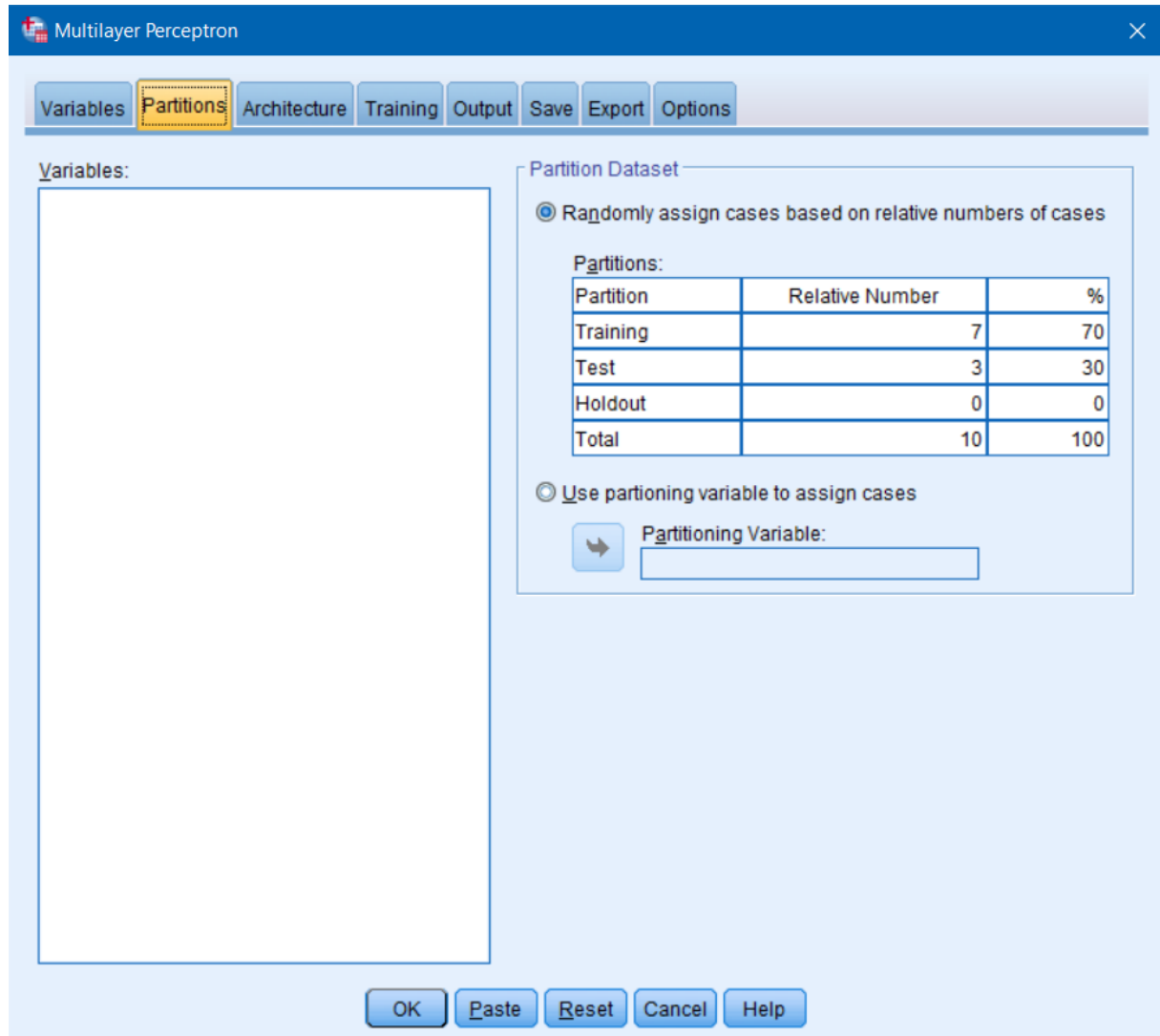


Figure 3.14 Setting Partitions

In the next step, detailed information about the architecture of the network is introduced into the software. Software offers two options to the user as automatic and custom architecture. In automatic selection user decides the maximum and minimum number of the nodes in a single hidden layer. However, in custom architecture mode more sophisticated parameter unlocked to the user namely number of hidden layers and number of nodes (neurons) in those hidden layers, type of the activation function as sigmoid or hyperbolic tangent and similarly type of activation function in output layer and re-scaling options such as normalized, adjusted normalized, standardized, etc. All these parameters have the potential to affect the performance of the

network positively or negatively. In this context, since there is no optimization or meta-heuristic approach in the software, the feasible result will be sought by trial-and-error method. The tab has been shaped as follows on the next page.

Multilayer Perceptron

Variables Partitions **Architecture** Training Output Save Export Options

☐ Automatic architecture selection

Minimum Number of Units in Hidden Layer: 7

Maximum Number of Units in Hidden Layer: 8

☒ Custom architecture

Hidden Layers

Number of Hidden Layers

☒ One

☐ Two

Activation Function

☐ Hyperbolic tangent

☒ Sigmoid

Number of Units

☐ Automatically compute

☒ Custom

Hidden Layer 1: 8

Hidden Layer 2:

Output Layer

Activation Function

☐ Identity

☐ Softmax

☐ Hyperbolic tangent

☒ Sigmoid

The activation function chosen for the output layer determines which rescaling methods are available.

Rescaling of Scale Dependent Variables

☐ Standardized

☒ Normalized

Correction: 0,02

☐ Adjusted Normalized

Correction: 0,02

☐ None

OK Paste Reset Cancel Help

Figure 3.15 Architectural Parameters

The next tab offers options in which the user can specify parameters related to network training. These parameters include the type of training and the general characteristics of the training. More specifically, the type of training menu user allowed us to choose the type of batch: online, mini-batch, or normal batch. In the training option parts, the software proposes two options: scaled conjugate gradient and gradient descent. Similarly, gradient descent selection software offers certain parameters, such as learning rate, momentum, and interval center and offset. Primarily, the values of these parameters were set as default, with rates of 0.4, 0.9, 0, and ± 0.5 , respectively. These and previous parameters were re-evaluated to improve the network

performance. The tab with the default selection forms is as follows:

Multilayer Perceptron

Variables Partitions Architecture **Training** Output Save Export Options

Type of Training

☒ Batch
☐ Online
☐ Mini-batch

Number of Records in Each Mini-batch

☒ Automatically compute
☐ Custom
 Number of Records:

Optimization Algorithm

☐ Scaled conjugate gradient
☒ Gradient descent

Training Options:

Option	Value
Initial Learning Rate	0,4
Momentum	0,9
Interval Center	0
Interval Offset	±0.5

OK Paste Reset Cancel Help

Figure 3.16 Training Parameters

The subsequent tab does not contain any parameters that may affect the network performance; however, it provides options and instruments that illustrate the performance of the network to the user, such as a detailed description of the network, network diagram, and synaptic weights for further usage. Moreover, model summary and charts that makes the output more comprehensive to analyze such as residual and predicted value charts have also been proposed to the user. Finally in this tab thanks to the independent variable importance analysis feature of the IBM SPSS software it is possible to see the relative importance of independent variables with respect to created forecasting model which allows user to analyze and understand the factors according to their importance. The tab forms as follows;

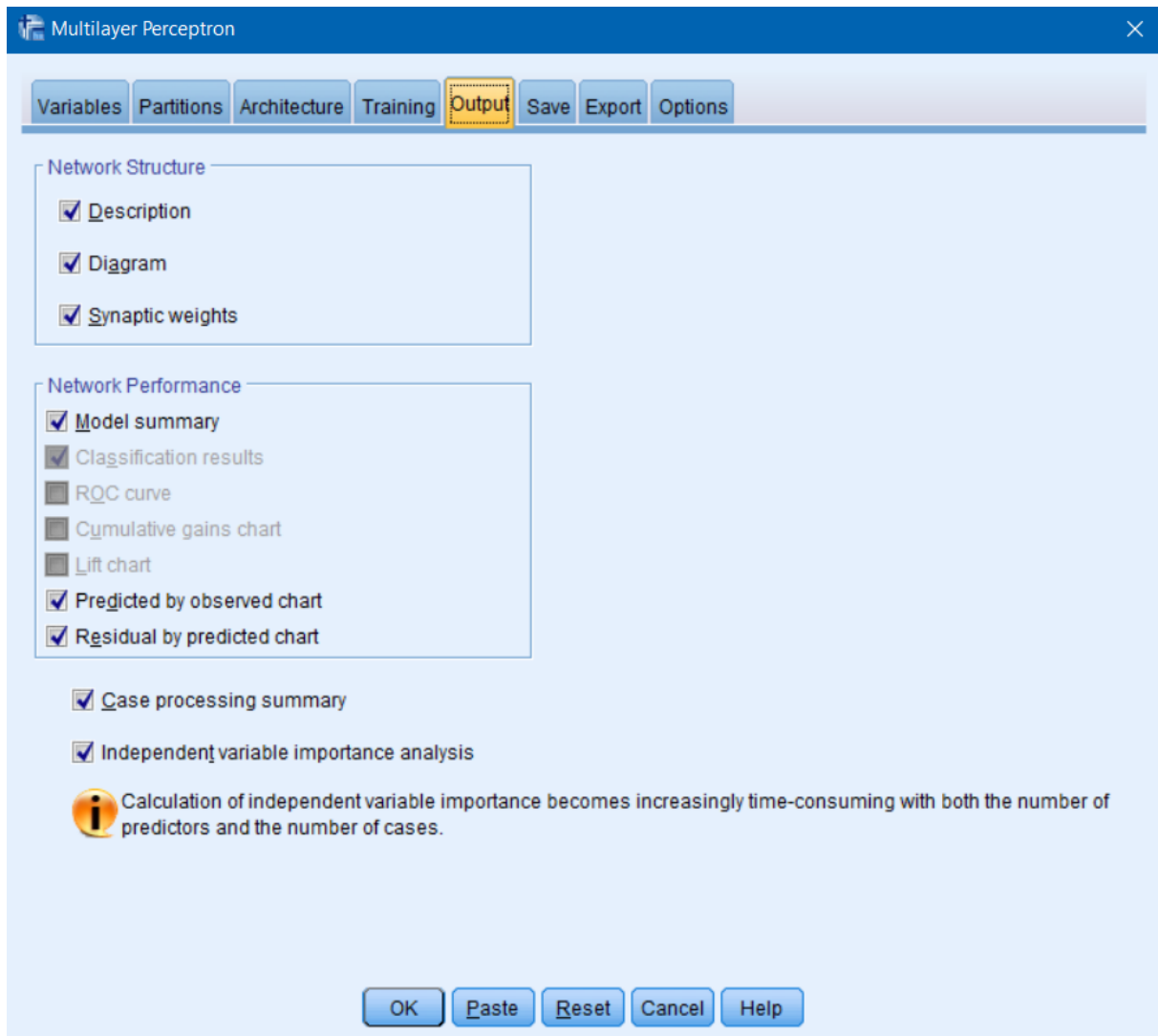


Figure 3.17 Output Options

The two tabs following this tab do not contain any parameters that affect the performance of the network. However, it provides convenience to the user in terms of creating a new column containing the forecasts that have been made and creating a new XLS file containing the synaptic weights for related model. As can be seen on the next page, a new column containing the forecasts is created as a result of activating related features in the software. Thanks to these new data, the user has the convenience of performing any action in the context of evaluating the performance of the network. In addition, recording the synaptic weights provides convenience to the user for the next prediction period if the model performance is appreciated. The tabs combined as follows;

Multilayer Perceptron

Variables | Partitions | Architecture | Training | Output | **Save** | Export | Options

☒ Save predicted value or category for each dependent variable
☐ Save predicted pseudo-probability for each dependent variable

Variables:

	Predicted Value or Category	Predicted Pseudo-Probability	
Dependent Variable	Name of Saved Variable	Root Name of Saved Variables	Categories to Save
sales	MLP_PredictedValue		25

Variables | Partitions | Architecture | Training | Output | **Save** | **Export** | Options

☐ Export synaptic weight estimates to XML file

Variables and File Names:

Dependent Variable	File Name
sales	Browse...

Names of Saved Variables

☒ Automatically generate unique names
 Select this option if you want to add a new set of saved variables to your dataset each time you run a model.

☐ Custom names
 Specify names for the variables. If you select this option, any existing variables with the same name or root name are replaced each time you run a model.

Figure 3.18 Saving and Exporting Options

In the last tab, the user has the option to specify the behavior of the software if some data are not available. In this study, there were no missing data. In addition, the software enquires to user when the training of the network should stop within the stopping rules framework. Among these, there are options such as the number of epochs that can affect the network performance. The number of epochs can be determined and proposed automatically as well as manually to the software. This part will be mentioned again in the section where the network performance improved but for primary action the values set as default as in previous parameter that have the potential of affecting the network performance both positively and negatively. In addition to these, various options such as how many cases the software will store and how many minutes the training can last have been presented to the user as follows;

Multilayer Perceptron

Variables Partitions Architecture Training Output Save Export **Options**

User-Missing Values
Specify how to treat cases with user-missing values on factors and categorical dependent variables.
☒ Exclude ☐ Include
 Cases with user-missing values on covariates or scale dependent variables are always excluded.

Stopping Rules
Stopping rules are tested in the order listed below.

Maximum steps without a decrease in error:

Data to Use for Computing Prediction Error:
☒ Choose automatically
☐ Both training and test data

☒ Maximum training time Minutes:

Maximum Training Epochs
☒ Compute automatically
☐ Specify custom value Maximum number of epochs:

Minimum Relative Change in Training Error:

Minimum Relative Change in Training Error Ratio:

Maximum Cases to Store in Memory:

OK Paste Reset Cancel Help

Figure 3.19 Other Options

After this stage, the artificial neural network was created, the network was trained using the default parameters, and the following relatively inaccurate results were obtained:

Table 3-3 Results with Default Parameters

MAD	4082.125
MSE	54614349.28
TS	-3.373488073
%MAPE	10.85113

As seen above, the MAPE value is 10.85%, which means that the model has an accuracy of 89.15%.

3.7.2. Improving the Performance of the Neural Network

As mentioned earlier, the software was primarily run using default parameters. However, many factors have the potential to affect the performance of the neural network, such as the consistency of the forecasts. Some of these factors are the number of data trained, number of hidden layers, number of nodes inside the hidden layers, type of activation function, learning rate, and momentum coefficient. Nevertheless, there are other parameters, such as learning algorithms, that cannot be interfered with but have the potential to affect the performance of the neural network because the software is a closed source. The parameters mentioned are considered default and will not be changed during the improvement process in this study.

3.7.2.1. Modifying Number of Hidden Layers

Because each artificial neural network is unique to its own dataset, it is equipped with different architectural elements. One of these elements is the number of hidden layers. However, it is rare to see more than two hidden layers in an artificial neural network, in addition to the input and output layers, unless a sophisticated deep learning strategy has been pre-determined and considered previously (Heaton, 2015). While the practical necessity of more than two hidden layers in most problems is unclear, it can be said that even two-hidden-layered artificial neural networks are seldom encountered in theory and in the literature as well. On the other hand, it would not be wrong to say that there are minor nuances between one and two hidden layer artificial neural networks. While the single hidden-layer neural network is capable of approximating any function that involves continuous congruency from a finite space to another a two-layer neural network is capable of embodying a random decision boundary to random accuracy with rational activation functions and may approximate any smooth congruency to any accuracy (Heaton, 2017). In this manner since there is no significant and special requirement for two or multiple hidden layered neural network architecture, in this study a single hidden layered ANN model have been decided to be utilized.

3.7.2.2. Modifying Number of Hidden Neurons

The number of hidden neurons may affect the performance of neural networks both positively and negatively. Thus, it is crucial to determine the accurate number of hidden neurons for efficient forecasting. Using too few or many hidden neurons in network architecture may occur under-fitting and over-fitting respectively. While, under-fitting may cause inadequate detection

of the signals from environment, over-fitting may induce overreach of the capacity of available data. Many methods have been proposed to determine the number of hidden neurons in network architectures that do not utilize any optimization approach. Some of those are, using the number of neurons between the sizes of input and output layers, using the number of neurons which are $2/3$ of input neuron and plus the size of output neuron and using the number of neurons less than twice of input layer (Heaton, 2017). Since in this research the size of the input layer is 9 and output layer is 1, for above three cases the number or ranges of hidden neurons could be calculated as

$$1 < n < 9, \quad (3.24.)$$

$$n = 7, \quad (3.25.)$$

$$\text{and } n < 18 \quad (3.26)$$

where n denotes the number of hidden neurons. As a result of these inferences, it was thought that the number of hidden neurons should be between 1 and 18, satisfying all the assumptions noted above. In the next step, only the number of hidden neurons was changed, and the other parameters were left as defaults.

Table 3-4 Results Improved with Number of Hidden Layers and Neurons

Number of Neurons	1	2	3	4	5	6	7
MAPE (%)	3.51839	4.13028	3.29879	3.20050	3.18863	3.55026	4.78228
Number of Neurons	8	9	10	11	12	13	14
MAPE (%)	3.99586	2.90301	3.86521	3.65912	6.99240	4.43996	3.75118
Number of Neurons	15	16	17	18			
MAPE (%)	3.22286	8.58631	3.99248	3.63319			

As a result of the experiments, the best MAPE value was found in a single layer neural network consisting of 9 hidden neurons with the degree of 2.90%. The closest value to this is a single-layer neural network consisting of 15 hidden neurons with a MAPE degree of 3.22%.

3.7.2.3. Modifying Initial Learning Rate

The initial learning rate, which is one of the training options of the software and is capable of affecting the performance of the neural network, is enabled by activating the gradient descent optimization algorithm section in the training tab. Correct determination of the learning rate is closely related to the performance of the neural network. The learning rate determines the amount of change in the weights, if an excessive value is assigned, it causes the neural network to circulate between the local solutions and oscillate. On the other hand, if a low value is assigned to learning rate it may cause an increment in training time. Although the learning rate is frequently used between 0.2 and 0.4 in the literature, it has been observed that a value of 0.6 produces efficient results in some models. In other words, the structure of the model and data set plays an important role in determining this value (Öztemel, 2020). More broadly, this coefficient ranges from 1.0 to any positive value greater than 0.0 depending on the type of problem, approach, and dataset and practiced in this manner.

As with many factors that can affect the performance of the artificial neural network, the learning rate will be tried to be determined by trial-and-error method in this study, just like the previous parameters. This process is realized by considering the number of hidden layers and the number of neurons in the hidden layers, which are among the previously determined parameters.

Other parameters that have not yet been determined were left at the default values determined by the software. In line with the information given above, the learning rate will be tested with trials between the values of 0.1 and 0.8 in terms of being comprehensive and the best value will be determined.

Table 3-5 Results Improved with Initial Learning Rate

Number of Neurons	9			
Learning Rate	0.1	0.2	0.3	0.4
MAPE (%)	8.13510	4.29179	3.98131	3.09520
Learning Rate	0.5	0.6	0.7	0.8
MAPE (%)	2.71409	4.51960	3.49184	2.88079

As shown above, the most accurate MAPE value was 2.71% when the learning rate was 0.5. The closest MAPE rate to this value was 2.88%, with a learning rate of 0.8. Thus, further improvement operations will be carried out by taking these values into consideration.

3.7.2.4. Modifying Momentum Coefficient

The momentum coefficient is another parameter that can positively or negatively affect the performance of a network. This can be defined as the addition of a certain ratio of the change in the previous iteration to the new change amount. This coefficient is a significant parameter that contributes to finding more accurate feasible results with a leap in networks which stuck to local solutions. Assigning very large numbers to this coefficient poses problems in reaching a single feasible solution, on the other hand assigning a very low value to this coefficient makes it difficult to escape from local solutions. Although experience shows that it would be appropriate to choose this value between 0.6 and 0.8, the nature of the problem and the type of data set is capable of disproof this statement. In this context, it would be best for the user to determine the momentum coefficient according to the data set and problem type on hand (Öztemel, 2020). Momentum coefficient is a very effective variable in terms of network performance. However, just like the learning rate, the momentum coefficient is not directly a structural element of neural networks. After the feasible result is achieved, the synaptic weights remain and the momentum coefficient is no longer needed (Heaton, 2015). In parallel with this information, it was decided to test the momentum coefficient between 0.1 and 0.9, taking the previously determined parameters into account.

Table 3-6 Results Improved with Momentum Coefficient

Number of Neurons	9			
Learning Rate	0.5			
Momentum Coefficient	0.1	0.2	0.3	0.4
MAPE (%)	4.28983	3.30253	2.54689	8.15079
Momentum Coefficient	0.5	0.6	0.7	0.8
MAPE (%)	3.34257	2.90096	3.44865	4.75916
Momentum Coefficient	0.9			
MAPE (%)	2.01368			
R²	0.9498			

From the results of the trials, it was determined that the best MAPE value of 2.01% was obtained when the momentum was 0.9. The closest value obtained with best MAPE has been determined on 0.3 momentum coefficient with 2.54% MAPE.

To rephrase this, the neural network achieved 97.99% accuracy with a momentum value of 0.9, a learning rate of 0.5, nine hidden neurons, and a single hidden layer. The table below summarizes the attributes of the artificial neural network together with its covariates (independent variables), number of hidden neurons, number of hidden layers, type of activation function for both input and output layers, and rescaling method for both dependent and independent variables, as follows:

Table 3-7 Model Summary

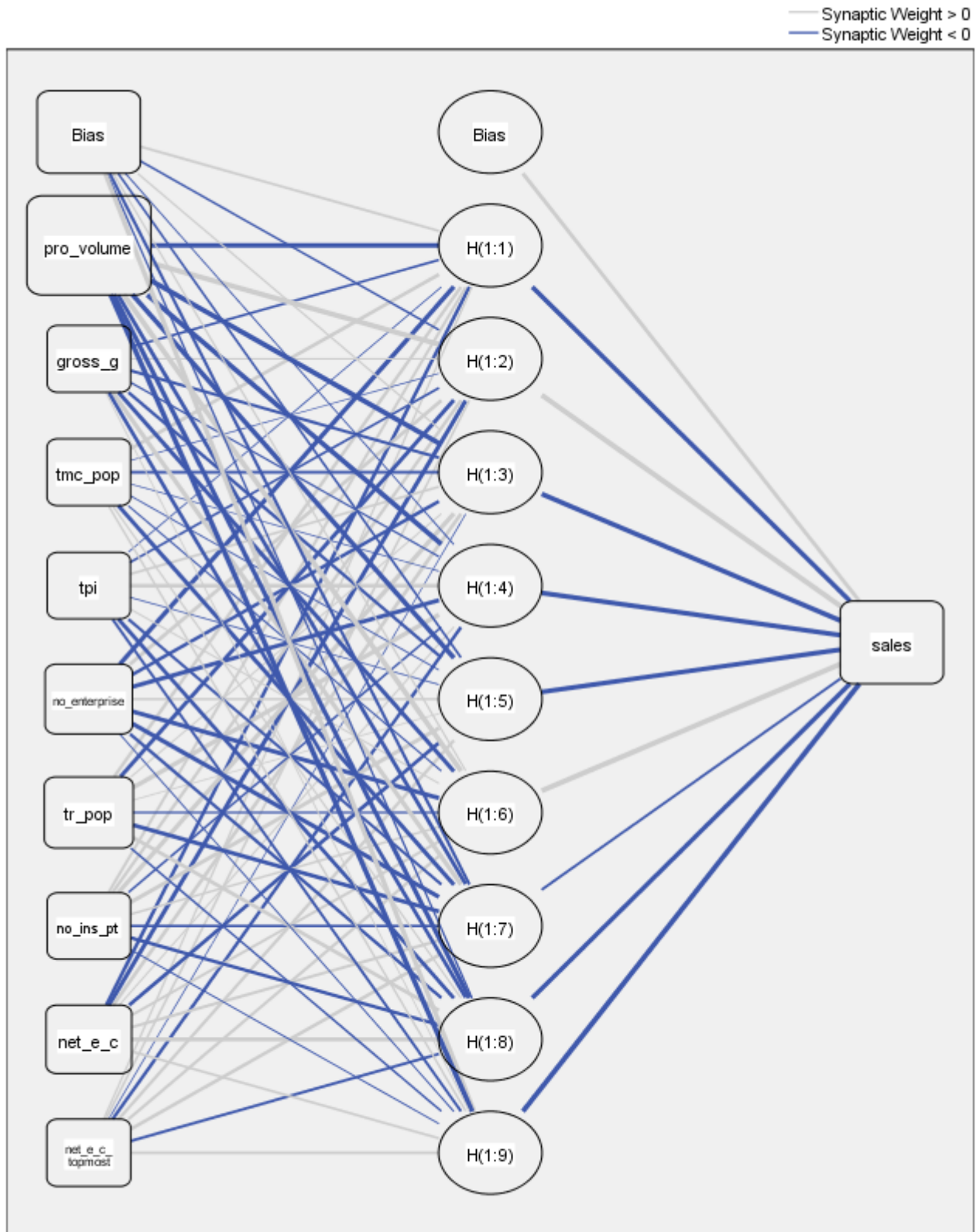
Network Information			
Input Layer	Covariates	1	pro_volume
		2	gross_g
		3	tmc_pop
		4	tpi
		5	no_enterprise
		6	tr_pop
		7	no_ins_pt
		8	net_e_c
		9	net_e_c_topmost
Hidden Layer(s)	Number of Units ^a		9
	Rescaling Method for Covariates		Normalized
	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1 ^a		9
Output Layer	Activation Function		Sigmoid
	Dependent Variables	1	sales
	Number of Units		1
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares

a. Excluding the bias unit

4. RESULTS

This study aims to forecast the demand for power transformers on country using artificial neural networks. Demand has been digitized by considering past sales data of power transformers. Factors affecting demand have been defined as production volume, population by country, total power installed, total power generated, electricity consumption, number of enterprises, electricity consumption of topmost exported countries, population of topmost exported countries, and the number of installed power transformers. Factors and dependent variables are introduced into IBM SPSS software, and the neural network toolbox of the software is utilized to forecast the demand for power transformers. Primarily, the toolbox runs with default parameter settings, and later on, parameters such as the number of hidden layers, number of hidden neurons in hidden layers, learning rate, and momentum coefficient are found via the trial-and-error method. Although the artificial neural network is capable of producing even more accurate outcomes with existent dependent and independent variables, once the satisfying results achieved, forecasting operations have been ended. In this context, the number of hidden layers, number of hidden neurons in the hidden layers, learning rate, and momentum coefficient are calculated as 1, 9, 0.5, and 0.9, respectively. Additionally, the most satisfactory result is obtained with 2.01% MAPE and 97.99% accuracy. Although artificial neural networks have experienced fluctuating periods in terms of reliability throughout history, as stated in the literature and theory, they are a high-potential forecasting tool that can produce consistent results. The highly consistent results revealed in this study once again revealed this situation. It is possible to note that this research also produced consistent outcomes with the literature review in the second chapter.

In this section, the artificial neural network model is presented along with the synaptic weights of the connections. In addition, parameter estimates are provided, which mathematically summarize the weights between hidden neurons. Furthermore, charts, such as the predicted value and residual, are proposed to illustrate the accuracy of the model. Finally, the relative importance scores of production volume, population by country, total power installed, total power generated, electricity consumption, number of enterprises, electricity consumption of topmost exported countries, population of topmost exported countries, and number of installed power transformers, together with their normalized importance charts, will be given.



Hidden layer activation function: Sigmoid

Output layer activation function: Sigmoid

Figure 4.1 Final Network Diagram

The above is a diagram of the artificial neural network. The connections between the hidden neurons and the input and output layers express synaptic weights, and the thicker the line, the more relevant the connection.

Table 4-1 Parameter Estimates

Parameter Estimates											
Predictor		Predicted									
		Hidden Layer 1									Output Layer
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)	H(1:9)	sales
Input Layer	(Bias)	,176	-,129	,082	-,047	-,141	,081	-,251	-,110	,952	
	pro_volume	-1,247	2,028	-1,722	-1,718	-1,347	1,968	-,385	-,774	-1,865	
	gross_g	-,152	,074	-,250	-,106	-,259	-,460	,508	-,364	,043	
	tmc_pop	,310	-,006	-,242	-,014	-,045	-,070	-,469	,077	,126	
	tpi	-,038	-,145	,151	,449	-,028	,064	-,225	-,385	-,197	
	no_enterprise	-,712	,335	-,307	-,474	,141	-,503	-,485	-,271	-,052	
	tr_pop	,279	-,413	,059	,707	,026	-,118	-,458	,341	-,128	
	no_ins_pt	,207	,258	,545	-,131	,388	,139	-,170	-,289	-,051	
	net_e_c	-,314	-,476	,326	,064	-,359	,182	,233	,464	,162	
	net_e_c_topmost	,107	,140	-,021	-,285	,252	,248	,296	-,212	,188	
Hidden Layer 1	(Bias)										,385
	H(1:1)										-1,252
	H(1:2)										2,256
	H(1:3)										-1,607
	H(1:4)										-1,385
	H(1:5)										-1,275
	H(1:6)										2,323
	H(1:7)										-,177
	H(1:8)										-,797
	H(1:9)										-2,251

The chart above expresses the weights between the layers and neurons mathematically. As can be seen, while the production volume gives the highest values in many connections, the closest values are seen in the number of installed power transformers, the number of enterprises and the population of the country.

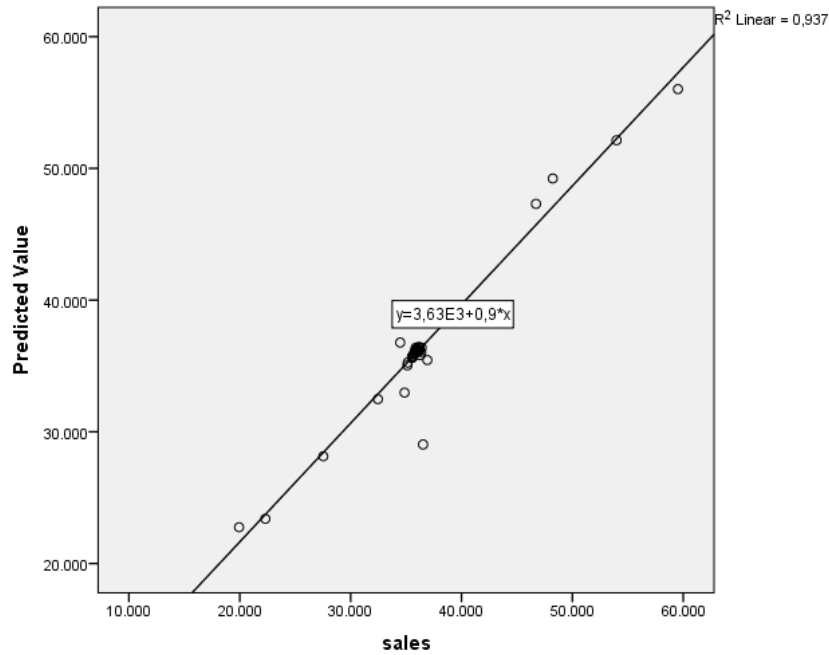


Figure 4.2 Sales-Predicted Value

Sales and predicted value graph, where the forecasted demand is plotted together with the fit line. As seen, thanks to the accuracy of the neural network model the plotted points have been placed regularly and in coherency with the fit line. The residual-by-predicted-value chart below shows no visible patterns between the residuals and predicted values.

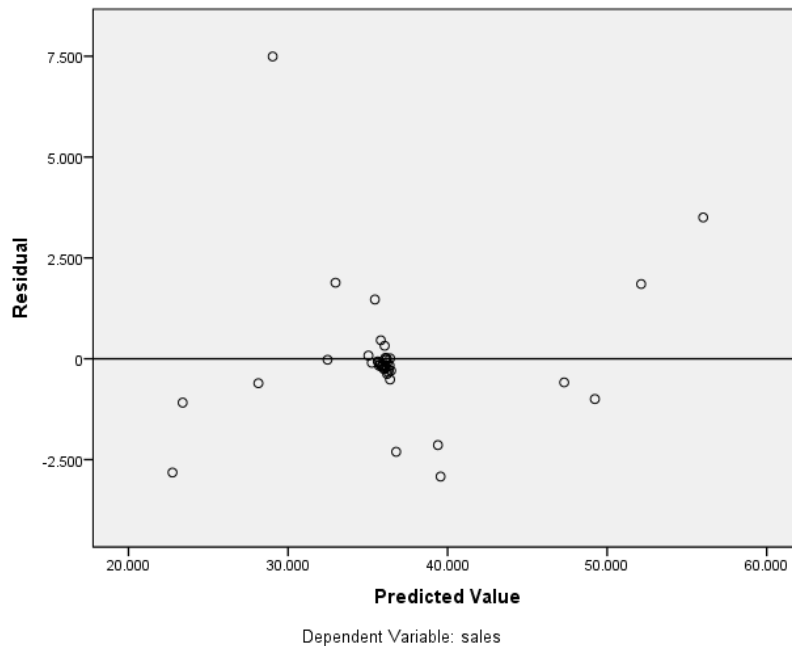


Figure 4.3 Predicted Value-Residual

Table 4-2 Independent Variable Importance Analysis

Independent Variable Importance		
	Importance	Normalized Importance
pro_volume	,700	100,0%
tr_pop	,011	1,6%
tpi	,016	2,3%
gross_g	,007	1,0%
net_e_c	,079	11,3%
no_ins_pt	,105	15,0%
no_enterprise	,010	1,4%
net_e_c_topmost	,050	7,2%
tmc_pop	,022	3,2%

Above table summarizes the independent importance of the variables in which pro_volume, tr_pop, tpi, gross_g, net_e_c, no_ins_pt, no_enterprise, net_e_c_topmost, and tmc_pop are production volume, population by country, total power installed, total power generated, electricity consumption, number of installed power transformers, number of enterprises, electricity consumption of the topmost exported countries, and population of the topmost exported countries, respectively. As it can be deduced, it has been determined that the most important factor is the volume of production with the importance level of ,700. The production volume is followed by population of country, number of enterprises and net electricity consumption of subjected country respectively. Net energy consumption of topmost exported countries and population of those come in the lower-middle ranks of the list, respectively, with values close to each other. The least important independent variable or factors in the estimation of total demand were gross generation, number of installed power transformers and total power installed respectively. In the bar-chart on the next page, the reflections of these values on the two-dimensional plane could be seen with respect to above values determined in model.

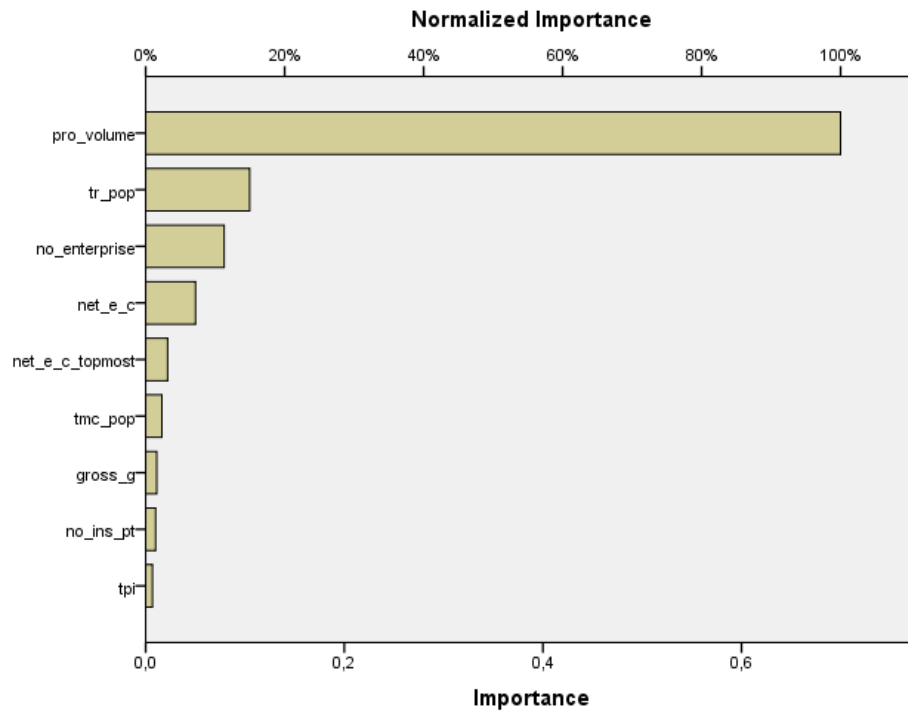


Figure 4.4 Normalized Importance Chart

The forecasts of the artificial neural network constructed could be seen as follows, together with the periods of each demand.

Table 4-3 Actual Demand vs Forecasted Demand

Period	Demand/Sales (unit)	Forecast (unit)
2021	53984.00	52131.00
2020	46718.00	47104.00
2019	36913.00	35444.00
2018	36375.00	36053.00
2017	27533.00	28139.00
2016	22307.00	22491.00
2015	19934.00	22754.00
2014	35147.00	35249.00
2013	34862.00	33975.00
2012	59520.00	58017.00
2011	37252.00	38391.00
2010	48233.00	49231.00
2009	35115.00	35036.00

2008	34469.00	36775.00
2007	32455.00	32479.00
2006	36639.00	38059.00
2005	36530.00	29036.00
2004	36401.00	36390.00
2003	36273.00	35813.00
2002	36202.00	36189.00
2001	36127.00	36102.00
2000	36198.00	36364.00
1999	36118.00	36144.00
1998	36157.00	36454.00
1997	36129.00	36224.00
1996	36047.00	36314.00
1995	35973.00	36188.00
1994	35899.00	36068.00
1993	35945.00	36278.00
1992	35875.00	36391.00
1991	35832.00	36211.00
1990	35825.00	36064.00
1989	35753.00	36009.00
1988	35750.00	35909.00
1987	35733.00	35901.00
1986	35666.00	35863.00
1985	35620.00	35759.00
1984	35594.00	35677.00
1983	35554.00	35619.00
1982	35526.00	35696.00

The performance of the model obtained by the error detection methods presented in the previous evaluation of the output section was calculated using several different methods, as shown in the table below.

Table 4-4 Performance Metrics of Neural Network

Period	Demand (sales)	Forecast	Error	Abs. Value	Squared Error	% Error
1982	35526.00	35696.00	-170.00	170	28900	0.478522772
1983	35554.00	35619.00	-65.00	65	4225	0.182820498
1984	35594.00	35677.00	-83.00	83	6889	0.233185368
1985	35620.00	35759.00	-139.00	139	19321	0.390230208
1986	35666.00	35863.00	-197.00	197	38809	0.552346773
1987	35733.00	35901.00	-168.00	168	28224	0.470153639
1988	35750.00	35909.00	-159.00	159	25281	0.444755245
1989	35753.00	36009.00	-256.00	256	65536	0.71602383
1990	35825.00	36064.00	-239.00	239	57121	0.667131891
1991	35832.00	36211.00	-379.00	379	143641	1.057713775
1992	35875.00	36391.00	-516.00	516	266256	1.438327526
1993	35945.00	36278.00	-333.00	333	110889	0.926415357
1994	35899.00	36068.00	-169.00	169	28561	0.470765202
1995	35973.00	36188.00	-215.00	215	46225	0.597670475
1996	36047.00	36314.00	-267.00	267	71289	0.740699642
1997	36129.00	36224.00	-95.00	95	9025	0.262946663
1998	36157.00	36454.00	-297.00	297	88209	0.821417706
1999	36118.00	36144.00	-26.00	26	676	0.071986267
2000	36198.00	36364.00	-166.00	166	27556	0.458588872
2001	36127.00	36102.00	25.00	25	625	0.069200321
2002	36202.00	36189.00	13.00	13	169	0.035909618
2003	36273.00	35813.00	460.00	460	211600	1.268160891
2004	36401.00	36390.00	11.00	11	121	0.03021895
2005	36530.00	29036.00	7494.00	7494	56160036	20.5146455
2006	36639.00	38059.00	-1420.00	1420	2016400	3.875651628
2007	32455.00	32479.00	-24.00	24	576	0.073948544
2008	34469.00	36775.00	-2306.00	2306	5317636	6.690069338
2009	35115.00	35036.00	79.00	79	6241	0.224975082
2010	48233.00	49231.00	-998.00	998	996004	2.0691228

2011	37252.00	38391.00	-1139.00	1139	1297321	3.057553957
2012	59520.00	58017.00	1503.00	1503	2259009	2.525201613
2013	34862.00	33975.00	887.00	887	786769	2.544317595
2014	35147.00	35249.00	-102.00	102	10404	0.290209691
2015	19934.00	22754.00	-2820.00	2820	7952400	14.14668406
2016	22307.00	22491.00	-184.00	184	33856	0.824853185
2017	27533.00	28139.00	-606.00	606	367236	2.200995169
2018	36375.00	36053.00	322.00	322	103684	0.885223368
2019	36913.00	35444.00	1469.00	1469	2157961	3.979627773
2020	46718.00	47104.00	-386.00	386	148996	0.826234
2021	53984.00	52131.00	1853.00	1853	3433609	3.432498518
Total						80.54700331

MAD	701
MSE	2108182.15
TS	0.273894437
%MAPE	2.01368
R ²	0.9498

As previously stated, the MAPE value was calculated to be 2.01%, which is a relatively satisfactory score for this neural network model. In addition to MAPE, the mean absolute deviation, mean squared error, and tracking signal were calculated as 701, 2108182.15, and 0.27, respectively, which means that the artificial neural network has also revealed its potential in different metrics.

4.1. A Comparison

Undoubtedly, one of the best ways to compare numerical approaches is to compare the outputs of these approaches with each other. In this context, comparing artificial neural networks, which are a relatively contemporary approach, with conventional statistical methods would be an effective method to reveal the consistency of the model once again. In parallel, Auto-Regressive Integrated Moving-Average (ARIMA) model have also been operated via IBM SPSS software. The model introduced by Box Jenkins (Box et al., 1994) and an ARIMA (p, d, q) forecasting model could be formulated as;

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (4.1.)$$

In this study, an ARIMA model was run with default parameters, which yielded the following outcomes:

Table 4-5 ARIMA Model Results

Model Statistics										
Model	Number of Predictors	Model Fit statistics					Ljung-Box Q(18)			Number of Outliers
		MAPE	MAE	MaxAPE	MaxAE	Normalized BIC	Statistics	DF	Sig.	
sales-Model_1	9	2,036	715,849	16,173	5908,034	15,436	17,965	18	,458	0

As seen above even Although the ARIMA model produced a significantly accurate forecasting output with a MAPE value of 2.03% MAPE value the artificial neural network model outperformed the ARIMA model with a 0.02% difference. The following graph summarizes the performance of the two approaches: customized (improved) artificial neural network model and autoregressive integrated moving average forecasting model.

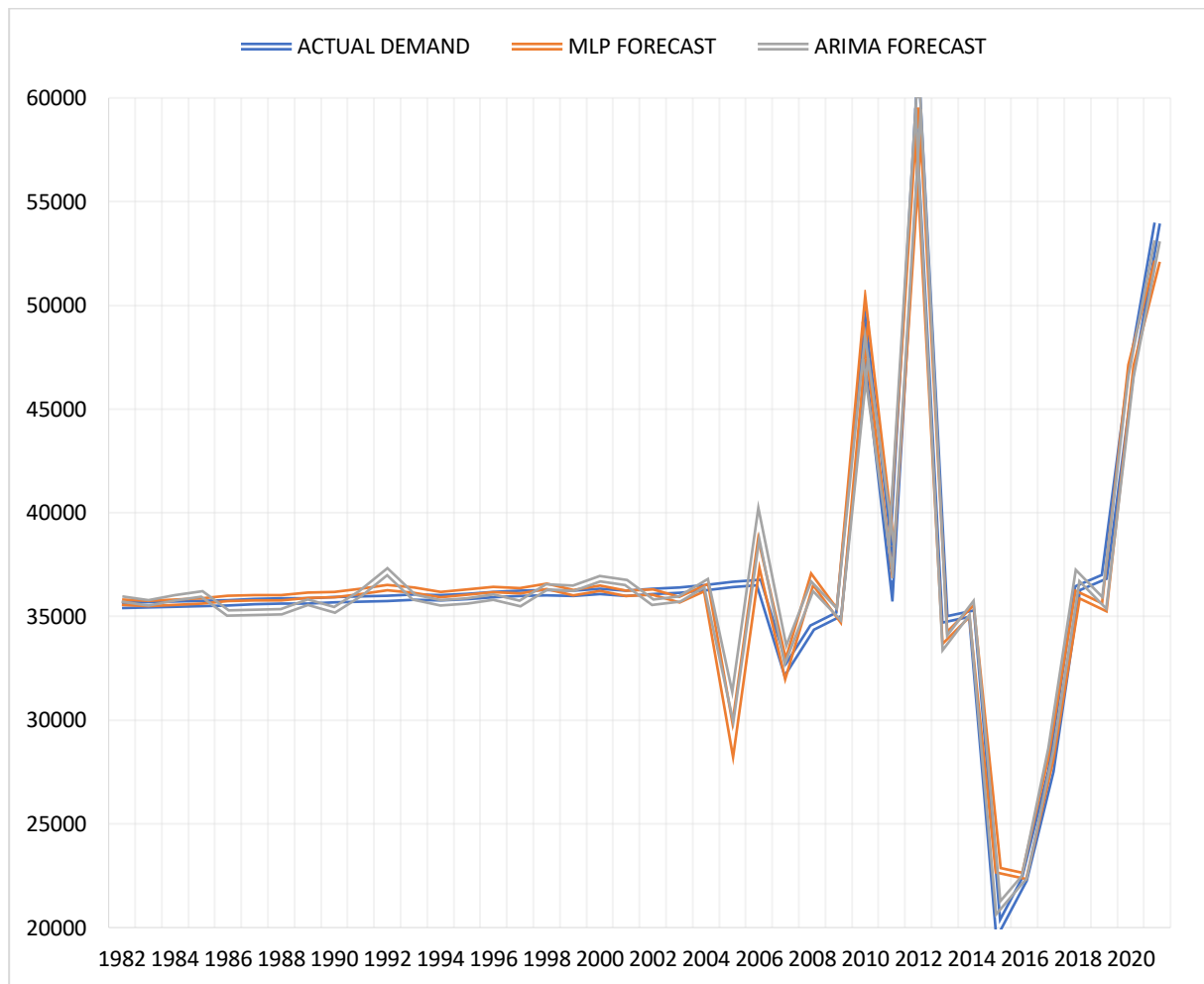


Figure 4.5 Comparison of Actual Demand with ANN and ARIMA Models

As seen in the legend, the colors blue, orange, and gray represent the actual sales, forecasted sales with the artificial neural networks, and sales forecasted by the ARIMA model, respectively. although ARIMA and ANN have produced results close to each other, as seen in the graph, ANN produced more accurate forecasts on most of the points (periods) and have been the one which approximate the actual sales data most.

5. CONCLUSION

In this study, demand forecasting for power transformers was performed using artificial neural networks. Factors that could potentially affect the demand for the product, namely, production volume, population by country, total power installed, total power generated, electricity consumption, number of enterprises, electricity consumption of the topmost exported countries, population of the topmost exported countries, and the number of installed power transformers, were decided by taking expert opinions. Additionally, historical sales data were gathered from official sources as a reference for product demand. After the compatibility issues of the data gathered from various sources are resolved artificial neural network application has been initiated.

A multilayer perceptron was utilized as the artificial neural network model. The procedure was supported by the IBM SPSS software. A three-layered network consisting of 1 output neuron, 9 input neurons and 9 hidden neurons, excluding the bias neuron, have been constructed via aforementioned software. Statistical approaches have been utilized to evaluate the performance of neural network created. As a result of several trials, satisfactory results have been achieved with a low error rate. Although the data were insufficient or incomplete for some factors, as a consequence of the modifications made in the infrastructure of the neural network, the performance of the neural network have been increased up to 97.99% accuracy.

Following consistency analysis, the factors affecting the performance of the network and the normalized importance results of the factors were investigated. Therefore, based on the outcomes achieved, it could be concluded that the hypothesis of 1) the demand for power transformers is driven by the production volume of the PTs population of the subject country and topmost exported countries, total power installed, total power generated, number of enterprises, electricity consumption of the topmost exported countries, and the country and number of installed PTs in the subject country to be strongly valid.

In addition to the ANN model, an autoregressive integrated moving-average (ARIMA) model was constructed for comparison. Notwithstanding the fact that the ARIMA model has also produced accurate outcomes according to calculations results it is possible to note that the multilayer perceptron artificial neural network model has outperformed the ARIMA model with the accuracies of 97.99% and 97.97% respectively.



6. RECOMMENDATIONS

As in this study, it is possible to say that the performance of artificial neural networks using a supervised learning strategy is directly proportional to both the number and quality of the collected data. Although many factors were used in this research, the rarity or lack of data held by official authorities forced the study in terms of the dataset used. It is undeniable that data are an important asset in today's world. In this sense, it is important for all small, medium or large-scale public or private institutions to keep data within the framework of legal limits. Nevertheless, in this study, the neural networks have produced forecasts with exceptional consistency thanks to their high potential and predictive ability. However, it is possible to obtain even better results with a larger data set or with different independent variables (factors). It could also be mentioned that eliminating factors that have fewer effect on performance of model and re-implement the process may suggested for future studies.

Another point to be discussed is optimization. As mentioned in the methodology section, artificial neural networks produce different results in each iteration. Some of these results may be extremely inconsistent, while others are considered sufficient by the network designer. Thus, the network designer terminates the approach as soon as it achieves a relatively satisfactory result. However, this may lead better local solutions and perhaps even global optimum solutions to go unnoticed. In this sense, in future studies, the inclusion of optimization approaches, such as particle swarm optimization and genetic algorithms, would be more reasonable.

Finally, it should be mentioned that there are no certain rules in the creation of artificial neural networks, determining the model, or determining the network architecture. Although such factors prevent the practical use of this approach in critical areas, it brings this approach to the fore as it models and solves problems that are impossible or very difficult to model mathematically. Users' expertise in the methodology may also reduce these problems to some extent.

REFERENCES

- Abd Rahman, N. H., Lee, M. H., Latif, M. T., & Suhartono. (2013). Forecasting of Air Pollution Index with artificial neural network. *Jurnal Teknologi*, 63(2).
- Abdel-Aal, R. E. (2008). Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks. *Computers & Industrial Engineering*, 54(4), 903–917.
- Adewole, A. P., Akinwale, A. T., & Akintomide, A. B. (2011). Artificial neural network model for forecasting foreign exchange rate. *World of Computer Science and Information Technology Journal (WCSIT)*, 1(3), 110–118.
- Afana, M., Ahmed, J., Harb, B., Abu-Nasser, B. S., & Abu-Naser, S. S. (2018). Artificial neural network for forecasting car mileage per gallon in the city. *International Journal of Advanced Science and Technology*. 124: 51-59.
- Aksoy, A., Ozturk, N., & Sucky, E. (2012). A decision support system for demand forecasting in the clothing industry. *International Journal of Clothing Science and Technology*, 24(4), 221–236.
- Aksoy, H., & Dahamsheh, A. (2009). Artificial neural network models for forecasting monthly precipitation in Jordan. *Stochastic Environmental Research and Risk Assessment: Research Journal*, 23(7), 917–931.
- Aktepe, A., Yanık, E., & Ersöz, S. (2021). Demand forecasting application with regression and artificial intelligence methods in a construction machinery company. *Journal of Intelligent Manufacturing*, 32(6), 1587-1604.
- Almonacid, F., Pérez-Higueras, P. J., Fernández, E. F., & Hontoria, L. (2014). A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. *Energy Conversion and Management*, 85, 389–398.
- Alonzo, R. J. (2010). *Transformers, capacitors and reactors, in: Electrical Codes, Standards, Recommended Practices and Regulations*, Elsevier, Burlington, MA. USA.
- Alvisi, S., Franchini, M., & Marinelli, A. (2007). A short-term, pattern-based model for water-demand forecasting. *Journal of Hydroinformatics*, 9(1), 39–50.

- Anthony, M., & Bartlett, P. L. (2011). *Neural network learning: Theoretical foundations*. Cambridge, England: Cambridge University Press.
- Armstrong, J. S. (Ed.). (2002). *Principles of forecasting: A handbook for researchers and practitioners* (2001st ed.). Dordrecht, Netherlands: Springer.
- Azadeh, A., Ghaderi, S. F., & Sohrabkhani, S. (2008). Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. *Energy Conversion and Management*, 49(8), 2272–2278.
- Baboo, S. S., & Shereef, I. K. (2010). An efficient weather forecasting system using artificial neural network. *International Journal of Environmental Science and Development*, 321–326.
- Barbosa, N. D. P., Christo, E. D. S., & Costa, K. A. (2015). Demand forecasting for production planning in a food company. *ARPJ Journal of Engineering and Applied Sciences*, 10(16), 7137-7141.
- Bashi, S. M., Abdullahi, U.U., Yunus, R., & Nordin, A. (2006). Use of natural vegetable oils as alternative dielectric transformer coolants. *Journal of the Institution of Engineers*.
- Bougadis, J., Adamowski, K., & Diduch, R. (2005). Short-term municipal water demand forecasting. *Hydrological Processes*, 19(1), 137–148.
- Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*. Prentice Hall, Englewood Cliffs, N.J.
- Braspenning, P. J., Thuijsman, F., & Weijters, A. J. (Eds.). (1995). *Artificial neural networks: An introduction to ANN theory and practice* (1995th ed.). Berlin, Germany: Springer.
- Buhari, M., & Adamu, S. S. (2012). Short-term load forecasting using artificial neural network. *Proceedings of the International Multi Conference of Engineers and Computer Scientists*. 1, 1-71.
- Burger, C. J. S. C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting — a case study of Durban, South Africa. *Tourism Management*, 22(4), 403–409.

- Campolo, M., Soldati, A., & Andreussi, P. (2003). Artificial neural network approach to flood forecasting in the River Arno. *Journal Des Sciences Hydrologiques [Hydrological Sciences Journal]*, 48(3), 381–398.
- Catalao, J. P. S., Pousinho, H. M. I., & Mendes, V. M. F. (2009). An artificial neural network approach for short-term wind power forecasting in Portugal. *2009 15th International Conference on Intelligent System Applications to Power Systems*, 1, 5–11. IEEE.
- Chae, Y. T., Horesh, R., Hwang, Y., & Lee, Y. M. (2016). Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings*, 111, 184–194.
- Chatfield, C. (2000). *Time-Series Forecasting*. Philadelphia, PA: Chapman & Hall/CRC.
- Chen, K.-Y., & Wang, C.-H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, 28(1), 215–226.
- Constantino, H. A., Fernandes, P. O., & Teixeira, J. P. (2016). Tourism demand modelling and forecasting with artificial neural network models: The Mozambique case study. *Tékhne*, 14(2), 113–124.
- Çuhadar, D. M., Cogurcu, I., & Kukrer, C. (2014). Modelling and forecasting cruise tourism demand to Izmir by different artificial neural network architectures. *International Journal of Business and Social Research*, 4(3), 12–28.
- Çunkaş, M., & Altun, A. A. (2010). Long term electricity demand forecasting in turkey using artificial neural networks. *Energy Sources Part B: Economics, Planning and Policy*, 5(3), 279–289.
- Demircioğlu, M., & Eşİyok, S. (2022). Energy consumption forecast of Turkey using artificial neural networks from a sustainability perspective. *International Journal of Sustainable Energy*, 1-15.
- Demirtaş, A. E., Anagün, A. (2007). Forecasting Method Selection for Capacity Planning in Service Systems. *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, 20 (1), 137-150.

- Dosdoğru, A. T., Boru, A., Göçken, M., Özçalıcı, M. & Göçken, T. (2018). Assessment of Hybrid Artificial Neural Networks and Metaheuristics for Stock Market Forecasting. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 27 (1), 63-78.
- Eğrioğlu, E., Yolcu., & U. Baş, E. (2019) *Yapay Sinir Ağları Öngörü ve Tahmin Uygulamaları*. Nobel Akademik Yayıncılık. Ankara.
- Enerdata. (2020). *Energy Reports: Iraq Energy Report*. Retrieved from <https://www.enerdata.net/estore/country-profiles/iraq.html>.
- Enerdata. (2020). *Germany Energy Information: Total Energy Consumption*. Retrieved from <https://www.enerdata.net/estore/energy-market/germany/>.
- Farimani, N. M., Parsafar, P., & Mohammadi, S. (2022). Evaluation performance of time series methods in demand forecasting: Box-Jenkins vs artificial neural network (Case study: Automotive Parts industry). *Journal of Statistical Computation and Simulation*, 1-20.
- Fente, D. N., & Kumar Singh, D. (2018). Weather forecasting using artificial neural network. 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). *IEEE*.
- Fernández, A., Ortiz, F., Delgado, C. Renedo., & S. Pérez. (2013). Comparative evaluation of alternative fluids for power transformers. *Electr. Power Syst. Res.* <http://dx.doi.org/10.1016/j.epsr.2013.01.007>.
- Fofana, I., Wasserberg, V., Borsi, H., & Gockenbach. E. (2002). Challenge of mixed insulating liquids for use in high-voltage transformers. Part 2: investigations of mixed liquid impregnated paper insulation, *IEEE Electrical Insulation Magazine*.
- Ghiassi, M., Lio, D., & Moon, B. (2015). Pre-production forecasting of movie revenues with a dynamic artificial neural network. *Expert Systems with Applications*, 42(6), 3176–3193.
- Ghiassi, M., Zimbra, D. K., & Saidane, H. (2008). Urban water demand forecasting with a dynamic artificial neural network model. *Journal of Water Resources Planning and Management*, 134(2), 138–146.
- Göçken, M., Boru, A., Dosdoğru, A. T., & Özçalıcı, M. (2018). Integrating metaheuristics and artificial neural network for weather forecasting. *International Journal of Global Warming*, 14(4), 440.

- Goh, B. (1998). Forecasting residential construction demand in Singapore: a comparative study of the accuracy of time series, regression and artificial neural network techniques. *Engineering Construction and Architectural Management*, 5(3), 261–275.
- Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23(5), 499–510.
- González-Romera, E., Jaramillo-Morán, M. A., & Carmona-Fernández, D. (2008). Monthly electric energy demand forecasting with neural networks and Fourier series. *Energy Conversion and Management*, 49(11), 3135–3142.
- Government of United Kingdom. (2020, July). *Historical electricity data: 1920 to 2019*. Retrieved from <https://www.gov.uk/government/statistical-data-sets/historical-electricity-data>.
- Guan, P., Huang, D.-S., & Zhou, B.-S. (2004). Forecasting model for the incidence of hepatitis A based on artificial neural network. *World Journal of Gastroenterology: WJG*, 10(24), 3579–3582.
- Gündüz, A. Y. (2016, October). *Makina Sektör Raporu. Çerkezköy Ticaret ve Sanayi Odası*. https://www.cerkezkoytso.org.tr/uploads/docs/cerkezkoy_tso_makine_sektor_raporu.pdf.
- Güven, İ., & Şimşir, F. (2020). Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ANN) and support vector machines (SVM) methods. *Computers & Industrial Engineering*, 147, 106678.
- Hagan, M. T., Demuth, H. B., & Beale, M. H. (2014). *Neural Network Design (2nd Edition)* (2nd ed.). Martin Hagan.
- Hamzaçebi, C., Es, H. A., & Çakmak, R. (2019). Forecasting of Turkey's monthly electricity demand by seasonal artificial neural network. *Neural Computing & Applications*, 31(7), 2217–2231.
- Heathcote, M.J. (1998). *The J&P Transformer Book* (12th ed.). New York, NY: Newnes.
- Heaton, J. (2007). *Introduction to neural networks with java*. Heaton Research.

- Heaton, J. (2015). *Artificial intelligence for humans, volume 3: Deep learning and neural networks*. North Charleston, SC: Createspace Independent Publishing Platform.
- Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York, USA: Wiley.
- Hill, T., Marquez, L., O'Connor, M., & Remus, W. (1994). Artificial neural network models for forecasting and decision making. *International Journal of Forecasting*, 10(1), 5–15.
- Hornby, A. S. (2000). *Oxford Advanced Learner's Dictionary* (6th ed.; A. S. Hornby, Ed.). London, England: Oxford University Press.
- Hu, M., Li, H., Song, H., Li, X., and Law, R. (2022). Tourism demand forecasting using tourist-generated online review data. *Tour. Manag.* 90, 104490.
- Huang, L., Xie, G., Zhao, W., Gu, Y., & Huang, Y. (2021). Regional logistics demand forecasting: a BP neural network approach. *Complex & Intelligent Systems*.
- Huber, J., Gossmann, A., & Stuckenschmidt, H. (2017). Cluster-based hierarchical demand forecasting for perishable goods. *Expert Systems with Applications*, 76, 140–151.
- Hyndman, R. J., & Athanasopoulos, G. (2013). *Forecasting: Principles and practice*. Otexts.
- IHME COVID-19 health service utilization forecasting team, & Murray, C. J. L. (2020). Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months (p. 2020.03.27.20043752). doi:10.1101/2020.03.27.20043752
- International Energy Agency. (2018). *Energy data by category, indicator, country or region: Iraq*. Retrieved from <https://www.iea.org/data-and-statistics/data-tables?country=IRAQ&energy=Electricity&year=2007>.
- International Energy Agency. (2018). *Energy data by category, indicator, country or region: Germany*. Retrieved from <https://www.iea.org/data-and-statistics/data-tables?country=GERMANY&energy=Electricity&year=2007>.
- International Energy Agency. (2019). *Energy data by category, indicator, country or region: France*. Retrieved from <https://www.eia.gov/opendata/qb.php?sdid=INTL.2-2-FRA-BKWH.A>.

- International Trade Center. (2020). [Bilateral trade between Turkey and Europe Product: 85 Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles 2020]. Retrieved April 4, 2021 from Trade Map database.
- Istituto Nazionale di Statistica. (2020). *Popolazione e Famiglie*. Retrieved from <https://www.istat.it/it/popolazione>.
- Jain, A., Varshney, A. K., & Joshi, U. C. (2001). Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks. *Water Resources Management*, 15(5), 299–321.
- Kandananond, K. (2011). Forecasting electricity demand in Thailand with an artificial neural network approach. *Energies*, 4(8), 1246–1257.
- Kaya, A., Kaya, G., & Çebi, F. (2022). Forecasting automobile sales in Turkey with artificial neural networks. In *Research Anthology on Artificial Neural Network Applications* (pp. 1478-1489). IGI Global.
- Khalidi, R., El Afia, A., Chiheb, R., & Faizi, R. (2017). Artificial neural network based approach for blood demand forecasting: Fez transfusion blood center case study. *Proceedings of the 2nd International Conference on Big Data, Cloud and Applications*. New York, NY, USA: ACM.
- Khotanzad, A., Afkhami-Rohani, R., & Maratukulam, D. (1998). ANNSTLF-Artificial Neural Network Short-Term Load Forecaster- generation three. *IEEE Transactions on Power Systems: A Publication of the Power Engineering Society*, 13(4), 1413–1422.
- Kiş, Ö. (2007). Streamflow forecasting using different artificial neural network algorithms. *Journal of Hydrologic Engineering*, 12(5), 532–539.
- Kon, S. C., & Turner, L. W. (2005). Neural network forecasting of tourism demand. *Tourism Economics*, 11(3), 301–328.
- Kou, J., Liu, J., Li, Q., Fang, W., Chen, Z., Liu, L., & Guan, T. (2013). Photovoltaic power forecasting based on artificial neural network and meteorological data. 2013 IEEE International Conference of IEEE Region 10 (TENCON 2013). *IEEE*.

- Kulkarniand, S.V., Khaparde, S.A. (2013). *Transformer Engineering- Design, Technology, and Diagnostics* (2nd ed.). USA: CRC press.
- Lam, K. C., Yu, C. Y., & Lam, K. Y. (2008). An artificial neural network and entropy model for residential property price forecasting in Hong Kong. *Journal of Property Research*, 25(4), 321–342.
- Lee, C., Song, H., & Mjelde, J. (2008). The forecasting of International Expo tourism using quantitative and qualitative techniques. *Tourism Management*. 29. 1084-1098.
- Lee, T. L., & Jeng, D. S. (2002). Application of artificial neural networks in tide-forecasting. *Ocean Engineering*, 29(9), 1003–1022.
- L'Institut National de la Statistique et des Études Économiques. (2020). *Population et structure de la population*. Retrieved from <https://www.insee.fr/fr/statistiques/series/103088458>.
- Liu, B., Fu, C., Bielefield, A., & Liu, Y. (2017). Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network. *Energies*, 10(10), 1453.
- Makridakis, S. (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*. 9(4), 527–529.
- Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. (1997). *Forecasting: Methods and Applications* (3rd ed.). Nashville, TN: John Wiley & Sons.
- Manobanda, A., Otero, P., & Granda, N. (2022). Electric Energy Demand Forecasting in an Oil Production Company Using Artificial Neural Networks. In *Latest Advances in Electrical Engineering, and Electronics*. Springer.
- Melikoglu, M. (2013). Vision 2023: Forecasting Turkey's natural gas demand between 2013 and 2030. *Renewable and Sustainable Energy Reviews*, 22, 393–400.
- NASA. (2020, April 30). *Is Time Travel Possible?*. <https://spaceplace.nasa.gov/time-travel/en/>.
- Nunes Da Silva, I., Spatti, D. H., Andrade Flauzino, R., Liboni, L. H. B., & Dos Reis Alves, S. F. (2018). *Artificial neural networks: A practical course*. Cham, Switzerland: Springer International Publishing.

- Office for National Statistics. (2020, June). *United Kingdom Population Mid-Year Estimate*. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/timeseries/ukpop/pop>.
- Öztemel, E. (2020). *Yapay Sinir Ağları*, Volume V. Papatya Yayıncılık Eğitim. Istanbul
- Pacelli, V., Bevilacqua, V., & Azzollini, M. (2011). An artificial neural network model to forecast exchange rates. *Journal of Intelligent Learning Systems and Applications*, 03(02), 57–69.
- Panapakidis, I. P., & Dagoumas, A. S. (2017). Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model. *Energy (Oxford, England)*, 118, 231–245.
- Papalexopoulos, A. D., Hao, S., & Peng, T.-M. (1994). An implementation of a neural network based load forecasting model for the EMS. *IEEE Transactions on Power Systems : A Publication of the Power Engineering Society*, 9(4), 1956–1962.
- Patro, S. G. K., & Sahu, K. K. (2015). Normalization: A Preprocessing Stage. *International Advanced Research Journal in Science, Engineering and Technology*, 20–22.
- Pino, R., Parreno, J., Gomez, A., & Priore, P. (2008). Forecasting next-day price of electricity in the Spanish energy market using artificial neural networks. *Engineering Applications of Artificial Intelligence*, 21(1), 53–62.
- Plett, G. (2015). *Battery Management Systems*, Volume I. Artech House, Norwood.
- Polebitski, A. S., & Palmer, R. N. (2010). Seasonal residential water demand forecasting for census tracts. *Journal of Water Resources Planning and Management*, 136(1), 27–36.
- Priyadarshi, R., Panigrahi, A., Routroy, S., & Garg, G. K. (2019). Demand forecasting at retail stage for selected vegetables: a performance analysis. *Journal of Modelling in Management*, 14(4), 1042–1063.
- Puri, V., & Kumar, N. (2021). Wind energy forecasting using artificial neural network in Himalayan region. *Modeling Earth Systems and Environment*.
- Republic of Turkish Ministry of Treasury and Finance (2016). *Elektrikli Makinalar ve Kablolar Sektörü: Sektör Raporları*. Retrieved from <http://www.antakyatso.org.tr/dokumanlar/2017%20yayinlar%20veriler/sekt%C3%B6r>

<el%20sorunlar/ALMANYA%20MAK%C4%B0NE-ELEKTR%C4%B0KL%C4%B0%20KABLOLAR-2017.pdf>.

- Rodrigues, F., Cardeira, C., & Calado, J. M. F. (2014). The daily and hourly energy consumption and load forecasting using artificial neural network method: A case study using a set of 93 households in Portugal. *Energy Procedia*, 62, 220–229.
- Romano, M., Liong, S.-Y., Vu, M. T., Zemskyy, P., Doan, C. D., Dao, M. H., & Tkalich, P. (2009). Artificial neural network for tsunami forecasting. *Journal of Asian Earth Sciences*, 36(1), 29–37.
- Rožanec, J. M., Kažič, B., Škrjanc, M., Fortuna, B., & Mladenčić, D. (2021). Automotive OEM Demand Forecasting: A Comparative Study of Forecasting Algorithms and Strategies. *Applied Sciences*, 11(15), 6787.
- Saglam, M., Spataru, C., & Karaman, O. A. (2022). Electricity Demand Forecasting with Use of Artificial Intelligence: The Case of Gokceada Island. *Energies*, 15(16), 5950.
- Sanayi Genel Müdürlüğü. (2020). *Elektrik-Elektronik Sektörü Raporu: Sektörel Raporlar ve Analizler Serisi*. Retrieved from [https://www.sanayi.gov.tr/assets/pdf/plan-program/ElektrikElektronikSektorRaporu\(2020\).pdf](https://www.sanayi.gov.tr/assets/pdf/plan-program/ElektrikElektronikSektorRaporu(2020).pdf).
- Sen, J., & Chaudhuri, T. D. (2017). A time series analysis-based forecasting framework for the Indian healthcare sector. Retrieved from <http://arxiv.org/abs/1705.01144>
- Singh, S., Hussain, S., & Bazaz, M. A. (2017). Short term load forecasting using artificial neural network. 2017 Fourth International Conference on Image Information Processing (ICIIP). *IEEE*.
- Singhal, D., & Swarup, K. S. (2011). Electricity price forecasting using artificial neural networks. *International Journal of Electrical Power & Energy Systems*, 33(3), 550–555.
- Somarathne, E. D. T., Wijayakulasooriya, J. V., & Karunasinghe, D. S. K. (2022). Application of artificial neural network for short term electricity demand forecasting. *KJMS*. 4,106-116.
- Souza, R. F., Wanke, P., & Correa, H. (2020). Demand forecasting in the beauty industry using fuzzy inference systems. *Journal of Modelling in Management*, 15(4), 1389–1417.

- Sözen, A., Akçayol, M. A., & Arcaklioğlu, E. (2006). Forecasting net energy consumption using artificial neural network. *Energy Sources Part B: Economics, Planning and Policy*, 1(2), 147–155.
- Statistisches Bundesamt. (2020). *Bevölkerung nach Geschlecht und Staatsbürgerschaft*. Retrieved from <https://www.destatis.de/EN/Themes/Society-Environment/Population/Current-Population/Tables/liste-current-population.html;jsessionid=7BB3475726205BDADF7945B270B24A54.live742>.
- Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, 16(2), 1223–1240.
- Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *The Journal of the Operational Research Society*, 54(8), 799–805.
- Terna SpA. (2019). *Lo storico dei dati statistici sull'energia elettrica e l'ultimo bilancio elettrico*. Retrieved from <https://www.terna.it/it/sistema-elettrico/statistiche/pubblicazioni-statistiche>.
- The World Bank. (2020). [World Population Prospects: 2019 Revision]. Retrieved April, 5 from The World Bank database.
- The World Bank. (2022). GDP (constant LCU) – Türkiye. Retrieved November, 9 from the World Bank database.
- Tirkeş, G., Güray, C., & Celebi, N. (2017). Demand forecasting: A comparison between the Holt-Winters, Trend Analysis and Decomposition models. *Tehnicky Vjesnik*, 24(2), 503–509.
- Tsao, Y.-C., Chen, Y.-K., Chiu, S.-H., Lu, J.-C., & Vu, T.-L. (2021). An innovative demand forecasting approach for the server industry. *Technovation*, 102371.
- Türkiye Elektrik İletim A.Ş. (2019). *Elektrik Enerjisi Üretimi-Tüketimi-Kayıplar 2019: Türkiye Brüt Elektrik Enerjisi Üretim-İthalat-İhracat ve Talebinin Yıllar İtibariyle Gelişimi (1996-2019)*. Retrieved from: <https://www.teias.gov.tr/tr-TR/turkiye-elektrik-uretim-iletim-istatistikleri>.

- Türkiye Elektrik İletim A.Ş. (2019). *Elektrik Enerjisi Üretimi-Tüketimi-Kayıplar 2019: Türkiye Elektrik Enerjisi Üretim Tüketim ve Kayıplarının Yıllar İtibariyle Gelişimi (1993-2019)*. Retrieved from: <https://www.teias.gov.tr/tr-TR/turkiye-elektrik-uretim-iletim-istatistikleri>.
- Türkiye Elektrik İletim A.Ş. (2019). *Enerji Nakil Hat Ve Trafoları 2019: Türkiye Trafo Adet ve Güçlerinin Primer Gerilimlerine Göre Yıllar İtibariyle Gelişimi (1980-2019)*. Retrieved from: <https://www.teias.gov.tr/tr-TR/turkiye-elektrik-uretim-iletim-istatistikleri>.
- Türkiye Elektrik İletim A.Ş. (2019). *Kurulu Güç: Türkiye Kurulu Gücünün Yıllar İtibariyle Gelişimi (1913-2019)*. Retrieved from: <https://www.teias.gov.tr/tr-TR/turkiye-elektrik-uretim-iletim-istatistikleri>.
- Türkiye Elektrik İletim A.Ş. (2022). *Yıllar İtibariyle Maksimum ve Minimum Günlük Tüketim Miktarları ve Tarihleri*. Retrieved from: <https://www.teias.gov.tr/turkiye-elektrik-uretim-iletim-istatistikleri>.
- Türkiye İhracatçıları Meclisi. (2019). *İhracat 2019 Raporu: Yeni Vizyon, Yeni Yol Haritası*. Retrieved from <https://www.immib.org.tr/files/kio/temp/Yeni%20Vizyon%20Yeni%20Yol%20Haritas%C4%B1%20%C4%B0hracat%202019%20Raporu.pdf>.
- Türkiye İstatistik Kurumu. (2019). *Yıllık Sanayi Ürün (PRODCOM) İstatistikleri, 2019: Toplam üretim ve satış değeri, 2010 - 2019*. Retrieved from <https://data.tuik.gov.tr/Kategori/GetKategori?p=Sanayi-114>.
- Türkiye İstatistik Kurumu. (2020). *Küçük ve Orta Büyüklükteki Girişim İstatistikleri: Ürün grubuna ve büyüklük grubuna göre ihracat*. Retrieved from <https://data.tuik.gov.tr/Kategori/GetKategori?p=sanayi-114&dil=1>.
- Türkiye İstatistik Kurumu. (2020). *Küçük ve Orta Büyüklükteki Girişim İstatistikleri: Ekonomik faaliyet ve büyüklük grubuna göre girişim sayısı*. Retrieved from <https://data.tuik.gov.tr/Search/Search?text=giri%C5%9Fim>.
- Türkiye İstatistik Kurumu. (2021). *Nüfus Büyüklüğü, Yaş ve Cinsiyet Yapısı: Yıllara, Yaş Grubu ve Cinsiyete Göre Nüfus, Genel Nüfus Sayımları – ADNKS*. Retrieved from <https://data.tuik.gov.tr/Search/Search?text=n%C3%BCfus>.

- Voronin, S., & Partanen, J. (2014). Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks: Forecasting electricity price and demand. *International Journal of Energy Research*, 38(5), 626–637.
- Wang, J., Liu, F., Song, Y., & Zhao, J. (2016). A novel model: Dynamic choice artificial neural network (DCANN) for an electricity price forecasting system. *Applied Soft Computing*, 48, 281–297.
- Wang, J.-M., & Liang, X.-H. (2009). The forecast of energy demand on artificial neural network. 2009 International Conference on Artificial Intelligence and Computational Intelligence. *IEEE*.
- Winders, J. J. Jr. (2002). *Power Transformers: Principles and Applications*. Power Engineering. New York, NY: Marcel Dekker, Inc.
- Yang, C.-L., & Sutrisno, H. (2018). Short-term sales forecast of perishable goods for franchise business. 2018 10th International Conference on Knowledge and Smart Technology (KST). *IEEE*.
- Yucesan, M., Gul, M., Mete, S., & Celik, E. (2019). A forecasting model for patient arrivals of an emergency department in healthcare management systems. In *Advances in Healthcare Information Systems and Administration* (pp. 266–284). *IGI Global*.

