T.C.

İSTANBUL KÜLTÜR UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

FORECASTING OF TURKEY'S TOTAL ELECTRICITY CONSUMPTION IN SECTORAL BASES USING MACHINE LEARNING ALGORITHMS

MASTER OF APPLIED SCIENCE THESIS

Mhd Khair HAJJAR

1900001489

Department: Industrial Engineering

Program: Engineering Management

Supervisor: Assist. Prof. İlayda ÜLKÜ

JUNE 2022

T.C.
İSTANBUL KÜLTÜR UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

FORECASTING OF TURKEY'S TOTAL ELECTRICITY CONSUMPTION IN SECTORAL BASES USING MACHINE LEARNING ALGORITHMS

MASTER OF APPLIED SCIENCE THESIS

Mhd Khair HAJJAR

1900001489

Department: Industrial Engineering

Program: Engineering Management

Supervisor: Assist. Prof. İlayda ÜLKÜ

Members of Examining Committee: Assist. Prof. Duygun Fatih DEMİREL

Assist. Prof. Mehmet ERDEM

JUNE 2022

ACKNOWLEDGEMENT

This research would not have been possible without the constant encouragement, support, patience, and guidance of my advisor supervisor Assist. Prof. İlayda Ülkü. I am most grateful for her invaluable assistance and insight. She has been the embodiment of a great leader and the ultimate role model.

I would also like to extend my gratitude to all the professors and doctors who have not hesitated to share their wealth of knowledge.

Last, but not least, my heartfelt thanks go to my family for their undoubted belief in me and overwhelming hope that empowered me to overcome any obstacle.

June 2022 Mhd Khair HAJJAR

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

ANN	Artificial neural network
APE	Absolute percentage error
ARFF	Attribute relation file format
ARIMA	Autoregressive integrated moving average
B	The bias
C	The penalty factor where it must be greater than zero.
ES	Exponential smoothing
GDP	Gross domestic production
GMDH	Group method data handling
GNP	Gross national production
GP	Gaussian process
GPR	Gaussian process regression
GWh	Giga watt hour
LR	Linear regression
LS-SVM	Least square support vector machine
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MLP	Multi-layer perceptron
MLPRegressor	Multi-layer perceptron regression
MSE	Mean squared error
NLR	Non-linear regression
$\mathbf{R^2}$	Determination coefficient
REPTree	Reduced error pruning tree
RMSE	Root mean square error

RS	Regression with seasonality
SMO	Sequential minimal optimization
SMOReg	Sequential minimal optimization for regression
SSE	Sum squared error
SVC	Support vector classification
SVM	Support vector machine
SVR	Support vector regression
SVRCIA	Seasonal support vector regression with chaotic immune algorithm
TEDAŞ	Electricity consumption and distribution statistics of Turkey
TÜİK	Turkish statistical institute
VAR	Vector autoregressive
WEKA	Waikato Environment for Knowledge Analysis
$\boldsymbol{w}$	Weight vector
$\boldsymbol{X}$	Input (actual) data
$\overline{\boldsymbol{x}_i}$	Mean of actual data
XNV	Correlated Nystrom views
$\boldsymbol{Y}$	Output (forecasted) data
$\boldsymbol{\varphi}$	Nonlinear activation function
$\boldsymbol{\xi_i, \xi_i^*}$	Nonnegative slack variables
$\boldsymbol{\varepsilon}$	Insensitive loss function

Üniversite	**: İstanbul Kültür Üniversitesi**
Enstitü	**: Lisansüstü Eğitim Enstitüsü**
Anabilim Dalı	**: Endüstri Mühendisliği**
Programı	**: Mühendislik Yönetimi**
Tez Danışmanı	**: Dr. Öğretim Üyesi İlayda ÜLKÜ**
Tez Türü ve Tarihi	**: Yüksek Lisans – June 2022**

ÖZET

Makine Öğrenimi Algoritmaları Kullanılarak Sektörel Bazda Türkiye'nin Toplam Elektrik Tüketiminin Tahmini

Mhd Khair HAJJAR

Türkiye, dünyanın en hızlı büyüyen ekonomilerinden birine sahip olarak kabul edilir. Elektrik enerjisi, her ülkenin ekonomik büyümesindeki kilometre taşlarından biridir. Bu nedenle, elektrik tüketimi hakkında bir fikir sahibi olmak oldukça önemlidir. Elektrik tüketimini tahmin etmek için literatürdeki çalışmalarda çeşitli makine öğrenme algoritmaları yöntemler kullanılmıştır. Bu çalışma, Türkiye'de 2050 yılına kadar sektörel ve toplam elektrik tüketimini tahmin etmektedir. Bu çalışmada, tahmin formüllerini oluşturmak için bir model olarak çok katmanlı algılayıcı (Multilayer perceptron MLP) ve sıralı minimum optimizasyon (sequential minimal optimization SMO) olmak üzere iki farklı zaman serisi tahmin yöntemi kullanılmıştır. 1970'den 2020'ye kadar Türkiye'nin sektörel ve toplam elektrik tüketimi Türkiye İstatistik Kurumu'ndan elde edilerek gelecek yılları tahmin etmek için modellerde girdi olarak kullanılmıştır. İki model, belirleme katsayısı (determination factor R^2) ve ortalama mutlak yüzde hatası (mean absolute percentage error MAPE) kullanılarak değerlendirilip ve karşılaştırılmıştır. Sıralı minimum optimizasyon regresyonu (SMOReg) modeli, R^2, MAPE ve ortalama kare kök hatası (root mean square error RMSE) için sırasıyla %91,42, %2,98 ve 8558,85 GWh olduğu Türkiye'nin toplam elektrik tüketimini tahmin etmede SMOReg'in sonuçları diğer modellerden daha üstün sonuçlar sunmuştur.

Sektörel bazda toplam elektrik tüketiminin 2021 yılında 292,142 GWh ve 2050 yılında 777,854 GWh'ye ulaştığı öngörülmüştür. SMOREg ayrıca sanayi ve ev tüketimi olarak sektörlerini tahmin etmede daha iyi performans göstermiştir. MLP’nin ise ticari, kamu, aydınlatma ve diğer sektör başlıklarını tahmin etmede daha iyi performans verdiği tespit edilmiştir.

Anahtar Kelimeler: Elektrik Tüketimi Tahmini, Makine Öğrenme, Yapay Sinir Ağı, WEKA, Çok Katmanlı Algılayıcı, SMO Regression.

University	**: İstanbul Kültür University**
Instittute	**: Institute of Graduate Studies**
Department	**: Industrial Engineering**
Program	**: Engineering Management**
Supervisor	**: Assist. Prof. İlayda ÜLKÜ**
Degree Awarded and Date	**: MS – June 2022**

ABSTRACT

Forecasting of Turkey's Total Electricity Consumption in Sectoral Bases Using Machine Learning Algorithms

Mhd Khair HAJJAR

Turkey is considered to have one of the fastest-growing economies in the world and electrical energy is a milestone in the economic growth of each country. Therefore, it is important to have an idea about the upcoming electricity consumption. Various methods were used in previous studies for forecasting electricity consumption. This study forecasts the sectoral and total electricity consumption in Turkey until the year 2050. This study utilizes two distinct Time series forecasting methods, namely Multilayer perceptron (MLP) and sequential minimal optimization (SMO) as a model to generate the forecasting formulas. The sectoral and total electricity consumption for Turkey from the year 1970 to 2020 was obtained from the Turkish Statistical Institute and fed to the models to forecast the upcoming years. The two models were evaluated and compared using determination coefficient R^2 and mean absolute percentage error (MAPE). It is found that SMOReg was superior in forecasting Turkey's total electricity consumption where the R^2, MAPE, and root mean square error (RMSE) for the SMOReg model were 91.42%, 2.98%, and 8558.85 GWh respectively, where the forecasted total electricity consumption in sectoral bases reaches 292,142 GWh in 2021 and 777,854 GWh in 2050. SMOReg also performed better in forecasting the industrial and household sectors. Whereas MLP performed better in forecasting the commercial, governmental, illumination, and other sectors.

Keywords: Forecasting electricity consumption, Machine learning, Artificial neural network, WEKA, Multilayer Perceptron, SMO Regression.

1. INTRODUCTION

Turkey's economy is considered one of the fastest-growing economies in the world and multiple sectors affect the economic growth according to the Organization for Economic Co-operation and Development (OECD, 2022). Wherein, the electricity consumption is deemed to have a significant factor. The forerunner to the economic development is the increase in the industrial production and demand along with the expansion in the residential areas. In which, the electricity consumption increases simultaneously with the growth of both sectors.

Various factors have an impact on electricity consumption, such as the gross domestic production (GDP), population, and revenue from industrial products. Many studies were conducted to forecast the electricity demand and consumption using different techniques such as artificial neural network (ANN), support vector machine (SVM), fuzzy logic, and hybrid models, such as forecasting electricity demand on seasonal bases in India (Usha and Balamurugan, 2016), forecasting energy consumption in Taiwan based on usage patterns (Chou and Tran, 2018) and estimating the impact of COVID-19 on the electricity demand in Turkey (Ceylan, 2021). These studies predict the electricity consumption through different periodicities which can be hourly, daily, monthly, or yearly.

This research aims to forecast total and sectoral electricity consumption for Turkey from the year 2021 until the year 2050 using the previous electricity consumption data for the period 1970-2020. MLP and SMO methods are used to create the forecasting formulas. An accurate forecast for electricity consumption in Turkey will be provided as this study will use evaluation metrics to obtain the best models.

This study is organized into five chapters. In the introduction part a brief summary of the study, its objectives, and its importance are provided. The following chapter shows Literature Review chapter with a historical background for forecasting electricity in Turkey and the world using different machine learning methods. Chapter 3 clarifies the Methodology chapter with the methods used in the study and their mathematical equations and the information about the used computer software. Implementation and Results chapter shows the used data, parameters, and numerical results in addition to evaluating and analyzing the results. Finally, the Conclusion part of this thesis presents a summarization of the study, key results, and recommendations concerning the findings of the study and future studies.

2. LITERATURE REVIEW

This literature presents studies about the utilization of different data and methods to forecast electricity load and consumption using machine learning algorithms and software. Abdel-Aal and Al-Garni (1997) developed an autoregressive integrated moving average (ARIMA) model to forecast the monthly electric energy consumption for one year using the past five years' data as an input where the average percentage error (APE) metric is used to evaluate the model. Saab, Badr, and Nasr (2001) forecasted Lebanon's electrical energy consumption using the autoregressive, ARIMA, and AR(1) with a highpass filter.

Kermanshahi and Iwamiya (2002) used a three-layered back-propagation and recurrent neural network (RNN) with a dataset for the period 1999-2000 to forecast the electrical load for nine power companies up until the year 2020 in Japan. In 2002 the cycle analysis gave better performance than other methods in forecasting energy demand in Turkey (Ediger & Useyin Tatlõdil, 2002). While Ceylan and Ozturk (2004) estimated the energy demand in Turkey until 2025 by developing a generic algorithm energy demand model where they used gross national production (GNP), population, and import and export figures from 1970 until 2001.

Yalcinoz and Eminoglu (2005) forecasted daily and monthly electricity consumption utilizing MLP using Matlab neural network toolbox with monthly data from 1991 to 2001. Tunç, Çamdli, and Parmaksizoglu (2006) used the electricity consumption for the years 1980-2001 to forecast the electricity consumption until the year 2020 with the constrained linear multivariable function in Matlab. Hamzaçebi (2007) also used Matlab to forecast the electricity consumption until the year 2020 in Turkey by developing an MLP model using the sectoral electricity consumption for the years 1970-2004 and used multiple statistical metrics to measure the model performance. Bilgili (2009) used linear regression (LR), nonlinear regression (NLR), and artificial neural network (ANN) to forecast electricity consumption in Turkey from 2008 until 2012.

Table 2.1 provides a summary of recent studies regarding forecasting electricity consumption and the algorithms used in this study alongside the parameters and the evaluation metrics.

Table 2.1 Review of recent studies on forecasting electricity consumption

Publisher(s)	Region	Used algorithms	Input factors	Periodicity	Dataset	Evaluation metrics
Hamzaçebi (2007)	Turkey	ANN	Electricity consumption sectors	Yearly	1970-2004	MAE, RMSE, APE, and MAPE
Bılgılı (2009)	Turkey	LR, NLR, and ANN	Installed capacity, gross electricity generation, population, total subscribership, and net electricity consumption	Yearly	1990-2007	relative error
Oğcu, Demirel, and Zaim (2012)	Turkey	ANN and SVM	Lag seasonal index, time index, and month index	Monthly	1970-2011	MAPE
Nwulu and Agboola (2012)	Cyprus	MLP	Historical electricity consumption, Gross domestic product (GDP), GDP per capita, and population	Yearly	1970-2002	RMSE and MSE
Kaytez, Taplamacioglu, Cam, and Hardalac (2015)	Turkey	MLP and Least Square Support Vector Machine (LS-SVM)	Gross electricity generation, installed capacity, total subscribership, and population	Yearly	1970-2009	MAPE, MSE, RMSE, SSE, and R^2
Usha and Balamurugan (2016)	India	Sequential Minimal Optimization Regression (SMOReg), MLP, Gaussin Regression (GR), and LR	Electricity consumption	Monthly	2008-2016	MAE, RMSE, and Direction accuracy
Deb, Zhang, Yang, Lee, and Shah (2017)	France	ANN, ARIMA, SVM, Fuzzy time series, case-based reasoning, gray, prediction moving average and exponential smoothing, and k-nearest neighbor hybrid models	Electricity load and climate data	Monthly and hourly	1996-2009	MAPE and RMSE
Ceylan (2020)	Turkey	Multiple Linear Regression (MLR), support vector regression (SVR), and GR	Agriculture value-added, total arable land, the GDP share of agriculture, and population	Yearly	1990-2017	MAPE, MAD, RMSE, and R2
Al Miraz, Biswas, and Rudra (2021)	Bangladesh	GMDH MLP	GDP, Gross National Income (GNI), CO2, ambient temp, and time series	Monthly	1995-2019	ME, MAE, MSE, RMSE, PE, and MAPE
Ceylan (2021)	Turkey	GP, SMOReg, correlated Nystrom views, LR, M5P model tree, and reduced error pruning tree	Daily load, curfew, temperature, national holiday, weekend, day of the week, actual time, and lagged variable	Daily	2019	MAPE, MAE, RMSE, and RRSE%
Comert and Yildiz (2021)	Turkey	ANN MLP	Temp, unemployment rate, and time series	Yearly and monthly	2000-2019	MAPE
Ağbulut (2022)	Turkey	Deep learning, SVM, ANN, and MLP	GDP per capita, population, vehicle kilometer, and year	Yearly	1970-2016	RMASE, MAPE, MBE, rRMSE, MABE, and R2
Velasquez, Zoctelli, Estanislau, and Castro (2022)	Brazil	ARIMA, exponential smoothing, and regression with seasonality	Average megawatt per month	Monthly	1999-2021	percent error
Ceylan, Akbulut, and Baytürk (2022)	Turkey	MLP, SMOreg, Random Forest, and GR	Daily gasoline consumption, day type (weekend or normal), workaday or holiday, curfew, and travel	Daily	2019-2020	MAPE, RMSPE, and R^2

Kandananond (2011) forecasted the electricity demand for Thailand using ARIMA, ANN, and MLP where the MAPE was used to evaluate the model's performance. Hong, Dong, Lai, Chen, and Wei (2011) found that seasonal support vector regression with chaotic immune algorithm (SVRCIA) performed better than ARIMA for forecasting monthly electricity load. Nwulu and Agboola (2012) used Matlab to train and develop an MLP model using data from 1993 until 2008 in order to forecast electricity consumption in north Cyprus.

Kaytez, Taplamacioglu, Cam, and Hardalac (2015) compared the performance of regression analysis, ANN, and LS-SVM in forecasting the electricity consumption in Turkey using Matlab, where LS-SVM is deemed to perform better and faster than the rest of the models as MAPE, mean square error (MSE), root mean square error (RMSE), sum square error (SSE), and determination coefficient R^2 were used to evaluate the models.

Becirovic and Cosovic (2016) used Matlab to create ANN model and Waikato Environment for Knowledge Analysis (WEKA) to generate a support vector machine for regression (SMOReg) model in order to forecast daily and hourly electricity consumption. Usha and Balamurugan (2016) forecasted monthly seasonal electricity consumption using SMOReg, linear regression (LR), Gaussian process regression (GPR), and MLP where WEKA is the software used to generate the model. After evaluating the models, SMOReg proved to be accurate in seasonal forecasting.

Deb, Zhang, Yang, Lee, and Shah (2017) generated and compared nine machine learning algorithms used to forecast building electricity consumption, the methods are ANN, SVM, ARIMA, Fuzzy times series, case-based reasoning, gray prediction, moving average, and exponential smoothing, K-nearest neighbor, and Hybrid models. The hybrid models are deemed to have the best performance as a different alternation of the models can be used to increase accuracy.

WEKA, Matlab, IBM SPSS, and RapidMiner studio are machine learning software that was used by Chou and Tran to generate ANN, SVM, LR, classification and regression tree and Hybrid models to forecast residential households' energy consumption where Hybrid models achieved the highest accuracy (Chou and Tran, 2018). Ceylan (2020) developed MLP, Bayesian SVR, and Bayesian GPR models to forecast Turkey's electricity consumption for the agriculture sector using Matlab. The

Bayesian GPR is deemed to have the highest accuracy. Bilal, Fayaz, Kim, and Pawar (2020) used various time series forecasting methods including MLP, K-nearest neighbor regression, SVR, LR, GPR, ARIMA, and vector autoregressive (VAR) to generate a model for forecasting the residential electricity consumption using WEKA and Python. SVR gave the best performance among the used methods. Al Miraz, Biswas, and Rudra (2021) utilized a group method of data handling (GMDH) using GMDH shell software to predict Bangladesh's electricity demand.

Ceylan (2021) utilized WEKA to forecast the electricity consumption during the curfew in Turkey. A Selection of machine learning algorithms was used including GPR, SMOReg, LR, correlated Nystrom views (XNV), reduced error pruning tree (REPTree), and M5P model tree. After performance benchmarking, SMOReg had the best performance.

Ağbulut (2022) forecasted the electricity consumption for the transportation sector and the CO2 emission where deep learning, SVM, and MLP were used to create the model. GDP per capita, population, vehicle kilometer, and year parameters from the year 1970 until 2016 were used in developing the models. SVM is considered to have the best accuracy according to a variety of statistical metrics used in evaluating the model performance.

Lydia and Kumar (2022) forecasted electricity demand by the use of LR, MLP, SMOReg, Baggings, and M5Rules in WEKA the generated models. M5Rules yielded the best performance after evaluating all the models. Comert and Yildiz (2021) used MLP to forecast the electricity consumption in Turkey, where MAPE evaluated the MLP-generated model.

Velasquez, Zoctelli, Estanislau, and Castro (2022) used Matlab to develop Regression with seasonality (RS), Exponential smoothing (ES), and ARIMA, which are time series methods. RS gave the best performance among them, and better performance can be obtained by combining the methods. Ceylan, Akbulut, and Baytürk (2022) used MLP, SMOReg, GPR, and Random Forest forecasting methods implemented in WEKA to forecast the gasoline consumption using daily gasoline consumption and other time-dependent variables as an input data. After evaluating the models Random Forest was deemed the most accurate.

3. METHODOLOGY

This section will present detailed information about where the data are obtained from, analysis of the data, forecasting software, machine learning algorithms, data preparation, and performance benchmark.

3.1. Data Collection

The data used in this paper consist of Turkey's total electricity consumption and the percentage of the electricity consumed by sectors from 1970 until 2020. The sectors are household, commercial, government, industrial, illumination, and others. Where other includes the consumption in agriculture, livestock, fishery sector, municipal water abstraction pumping facilities, and other public services. The data were obtained from the Turkish Statistical Institute (Türkiye İstatistik Kurumu) TÜİK (TurkStat - Data Portal, 2022) which were collected and reported from Electricity Consumption and Distribution Statistics of Turkey (TEDAŞ) (Geliştirme and Başkanliği, 2021).

3.2. Data Analyzation

The total electricity consumption in 1970 was 7308 GWh and through the years the consumption kept rising and reached 262702.13 GWh in 2020 with an average of 94481.06 GWh through that period as presented in Figure 3.1.

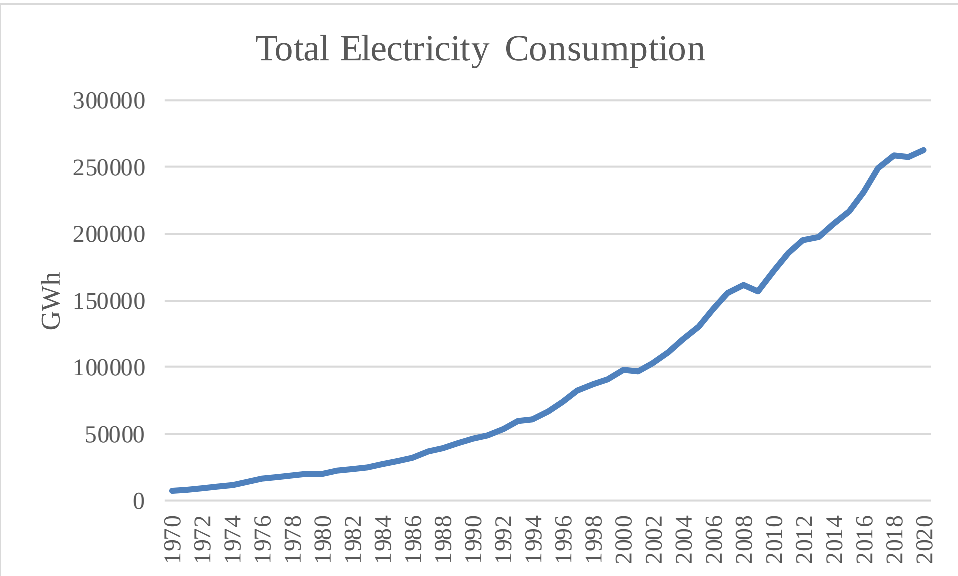

Figure 3.1 Total Electricity Consumption

Through the period of 51 years between 1970 and 2020, the industrial sector had the highest percentage of electricity consumption in Turkey. In 1970 the consumption percentage was 64.2% and in 2020 reached 45.69% with an average of 55.9% which accumulate to 4691.7 GWh and 120027.3 GWh respectively and an average of 46977.05 GWh as shown in Figure 3.2.

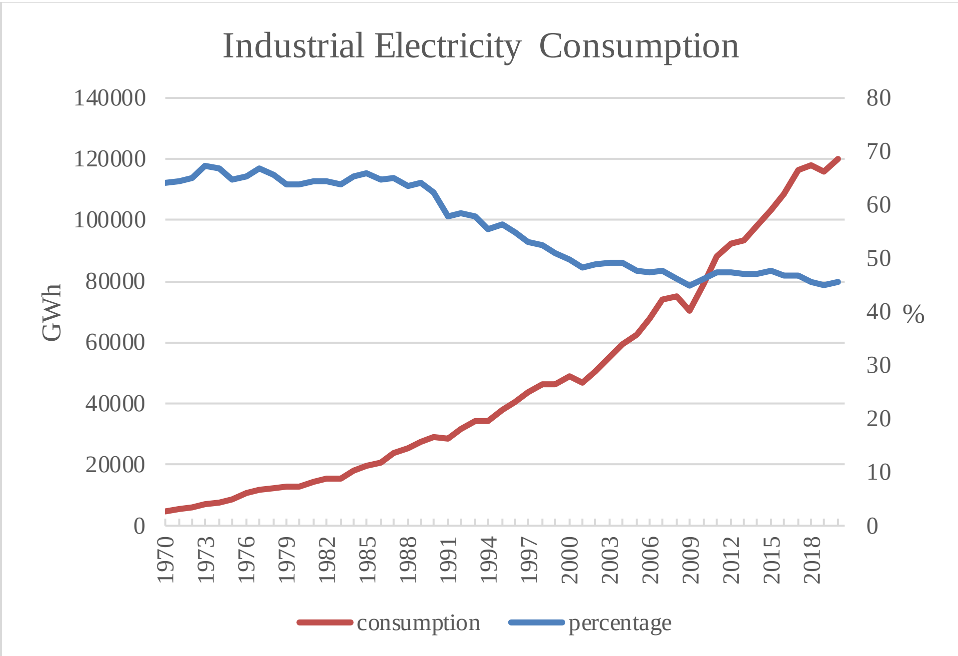

Figure 3.2 Industrial Electricity Consumption

The sector which came next was the household sector where the electricity consumption in 1970 was 15.9% and in 2020 was 23.1% of the total electricity consumption with an average of 21.1% where the electricity consumption was 1161.97 GWh and 60693.46 GWh respectively with an average of 21252.85 GWh as demonstrated in Figure 3.3.

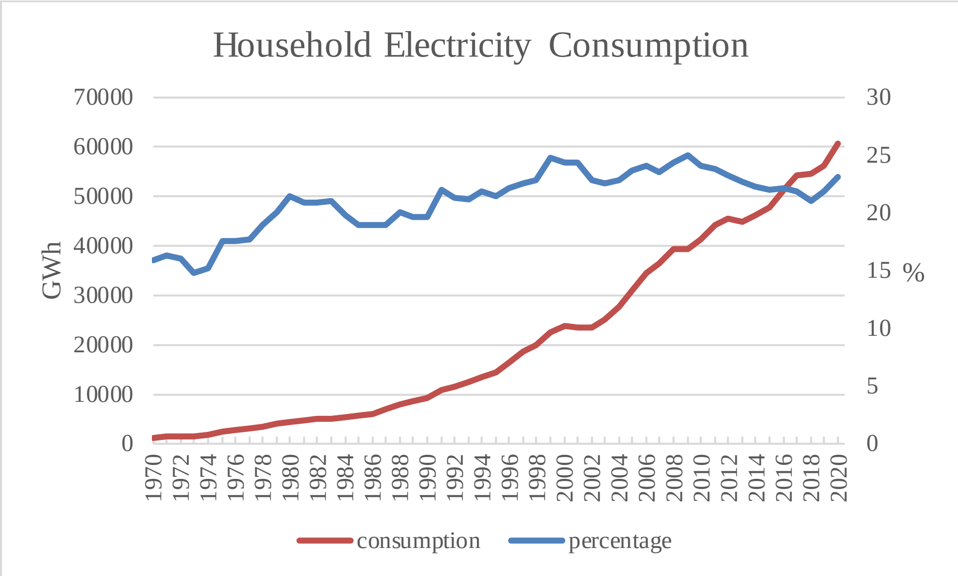

Figure 3.3 Household Electricity Consumption

The Commercial sector came after the household sector where the percentage of the electricity consumption was 4.8% in 1970 and reached 17.3% in 2020 with an average of 9.86% where the electricity consumption was 350.78 GWh and 45476.41 GWh respectively with an average of 13535.55 GWh as illustrated in Figure 3.4.

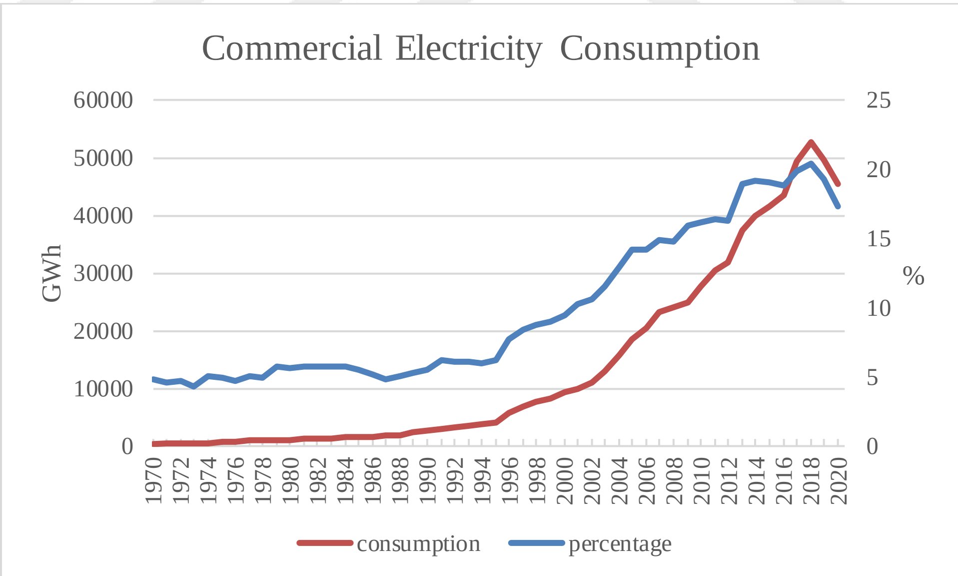

Figure 3.4 Commercial Electricity Consumption

Figure 3.5 represent The Governmental sector as it came afterword with electricity consumption of 299.6 GWh in 1970 and 12815.53 GWh in 2020 with an average of 3947.7 GWh during this period, where the share of the Governmental sector was 4.1% in 1970 and 4.87% in 2020 with an average of 3.85% of the total electricity consumption.

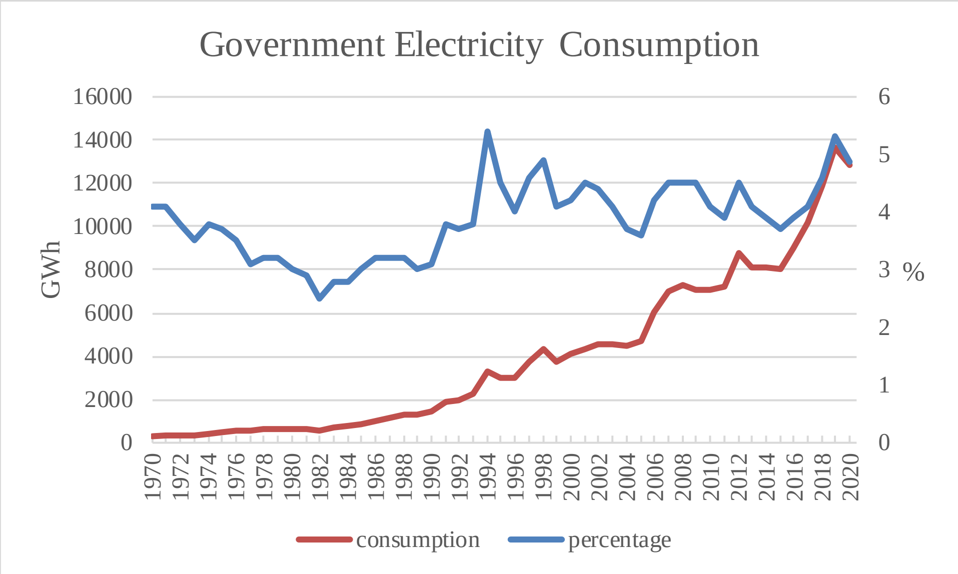

Figure 3.5 Government Electricity Consumption

After that came the Illumination sector with electricity consumption of 190 GWh in 1970 and 5237.5 GWh in 2020 with an average of 2458.4 GWh where the Illumination sector held 2.6% in 1970 and 1.99% in 2020 of the total electricity consumption during that time and the average percent was 2.58% as shown in Figure 3.6.

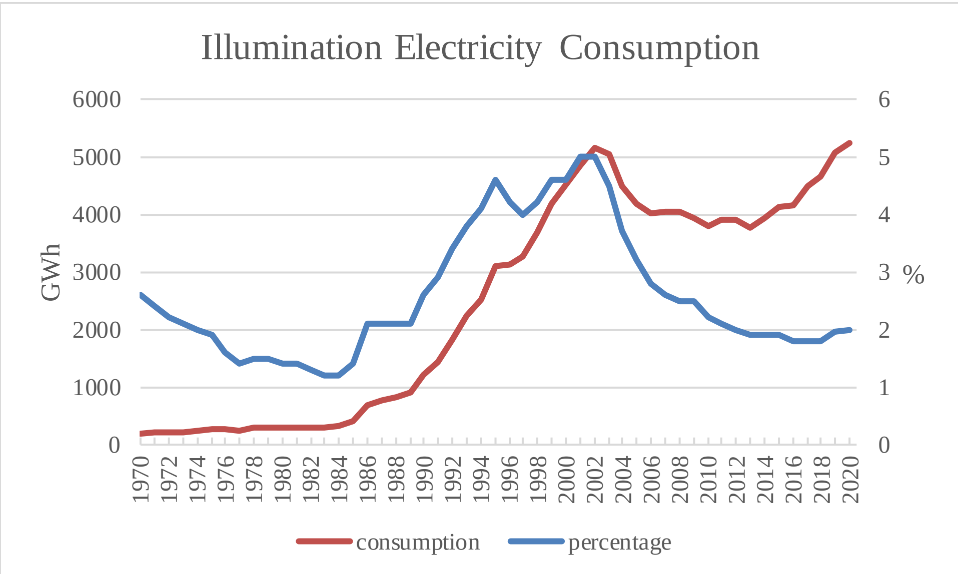

Figure 3.6 Illumination Electricity Consumption

The rest of the sectors were combined under the term Others which include the consumption in agriculture, livestock, fishery sector, municipal water abstraction pumping facilities, and other public services. The Other sectors consumed 613.8 GWh in 1970 with a share of 8.4% of the total electricity consumption and 262702.12 GWh in 2020 and the share was 7.02%, the average electricity consumption during that time was 94481.06 GWh and the average consumption percent was 6.68% illustrated in Figure 3.7.

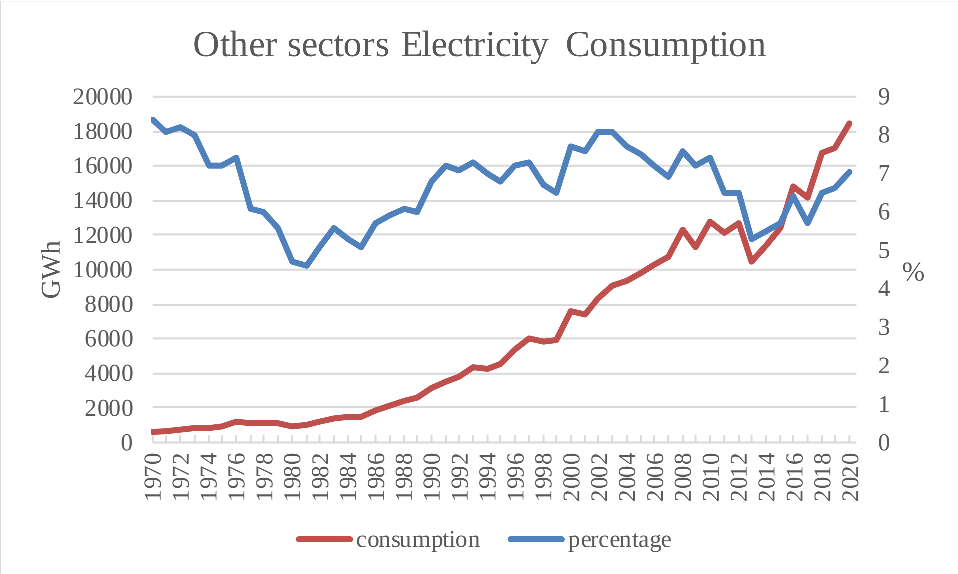

Figure 3.7 Other sectors Electricity Consumption

3.3. Forecasting Software

The software used to forecast the electricity consumption according to the previous data is WEKA (Weka 3 - Machine Learning Software in Java, n.d.). This software is based on Java and is considered a collection of Machine Learning algorithms for data mining as well as a lot of tools for classification, clustering, regression data preparation, and more.

WEKA is free and open-source software developed at the Waikato University of New Zealand. WEKA version 3.8.6 with the time series forecasting tool version 1.0.27 was used to generate the forecasting models using different machine learning models.

3.4. Machine Learning Models

WEKA contains multiple machine learning algorithms. In this study, Multi-Layer Perceptron Regressor (MLPRegressor) and Sequential Minimal Optimization for Regression (SMOReg) were applied. MLPRegressor is a model that employs regression as a base for the neural network which uses *partial_fit* parameter that enables to learn the models in real-time problems (Jain et al., 2021). SMOReg uses weights for each attribute to produce an output model that predicts the forecasted values (David et al., 2020).

3.4.1. Multi-Layer Perceptron

ANN is a simple method to emulate the human brain by using input, output, and hidden layers between them. A single layer consists of a number of nodes, each of which is connected with the nodes of the following layer by connections. A connection has a weight that is determined by training the ANN to attain a meaningful output from the input data.

MLP uses a nonlinear function and converts the input data into outputs using supervised learning algorithms. Given that the inputs are $X= x_1, x_2,\ldots,x_R$ where R is the total number of inputs and the outputs are $Y= y_1, y_2,\ldots,y_S$ where S is the total number of outputs, using the input vector X, the output vector Y can be calculated using the formula as seen in Equation (1):

$$y = \varphi(w^T x + b) \tag{1}$$

Where φ is the nonlinear activation function, w_i is the weights vector, and b is the bias (Feng et al., 2020). MLPRegressor is a specific type of MLP that is employed in forecasting problems (Csépe et al., 2014). There is no exact formula to determine the optimum parameters for the hidden neuron. For that, testing multiple numbers of neurons between the input layer and the output layer is the way to obtain a high-performance network.

3.4.2. Sequential Minimal Optimization for Regression

Sequential Minimal Optimization for Regression (SMOReg) is a machine learning algorithm based on a support vector machine (SVM) for regression.

SVM algorithm has a variety of applications. One of which is support vector classification (SVC) which is used for classification problems and the other is support vector regression (SVR) which is used to solve regression problems (Ceylan & Atalan, 2021).

SVM was presented in 1995 by Cortes and Vapnik with the concept of creating a hyperplane spatting into two classes where this hyperplane has the widest margin between them. Many hyperplanes can be created using SVR and that is where SVM comes to find the optimal hyperplane where the distance between classes has the maximum possible value (Cortes et al., 1995). SVM mathematical formula is demonstrated in Equation (2).

$$\text{minimize } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{2}$$

Where the conditions provided in Equations (3), (4), and (5) must be fulfilled the:

$$y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \tag{3}$$

$$\langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \tag{4}$$

$$\xi_i, \xi_i^* \geq 0 \tag{5}$$

w = the weights vector.

x = the input vector.

C = the penalty factor where it must be greater than zero.

ξ_i, ξ_i^*= the nonnegative slack variables.

ε = the insensitive loss function.

y = the target value.

Smola and Schölkopf evolved The Sequential Minimal Optimization (SMO) to solve regression problems using the SVM structure (Smola et al., 2004). SMOReg improved upon the calculation and data processing speed over SMO (Shevade et al., 2000).

Kernel function helps SVM to execute nonlinear transformation by reducing the complexity of the algorithm. This can be achieved by creating a map between the input and the high dimension space. The kernel used in this study is the polynomial kernel (Poly Kernel) according to the size of the used data and the mathematical model for this kernel is expressed in Equation (6).

$$K(x_i, x_j) = \left(x_i^T x_j + 1\right)^P \tag{6}$$

x_i, x_j are the inputs, x_i^T is the transpose of x_i and p is the degree of the polynomial equation (Taghi Sattari et al., 2021). Figure 3.8 shows the SVR slack variables and the hyperplane (Kavitha et al., 2017).

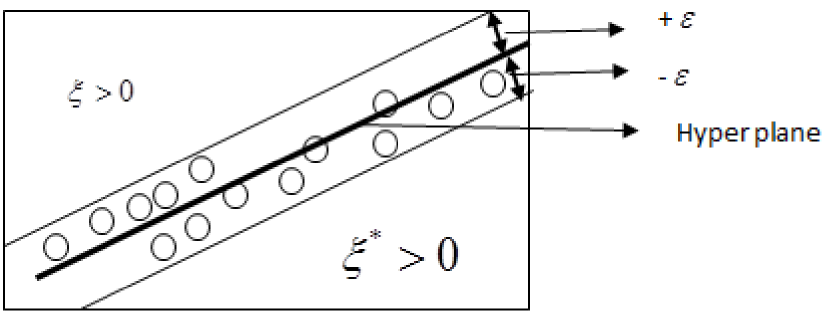

Figure 3.8 Support Vector Regression (Kavitha et al., 2017)

3.5. Performance Benchmark

Six statistical metrics were used to assess the performance of the machine learning algorithms used to forecast the results.

The metrics are R^2, MAE, MAPE, MSE, absolute percentage error (APE) and RMSE.

R^2 and MAPE are used to evaluate the model accuracy as well as performance, such that the parameters that yielded the highest R and the lowest MAPE without getting over or under fitting opted. Table 3.1 demonstrates the statistical metrics with the corresponding mathematical equations.

Table 3.1 Statistical performance metrics

Metric	Equation	Description	Performance Criteria
R^2	$1-\frac{\sum(y_i-x_i)^2}{\sum(x_i-\overline{x_i})^2}$	R^2 is a widely used metric that indicates how successful the model is in predicting the results and tracking the trends of the measured data.	The value of R^2 is between 0 and 1. When the value approaches one, better performance is indicated.
MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i-x_i\rvert$	MAE is used to determine the long-term performance of the model.	MAE is a positive value where the lower it is the better performance of the model is. Zero MAE means that the ideal model has been achieved.
MAPE	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{x_i-y_i}{x_i}\right\rvert\times 100$	MAPE calculates the average error for the predicted data in relation to the actual data.	The percentage of MAPE indicates the accuracy of the model. When its value is ≤10% the model is considered highly accurate, greater than 10%, and ≤20% the accuracy is considered good, greater than 20%, and ≤50% the accuracy is reasonable, higher than 50% the model is deemed inaccurate.
MSE	$\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i)^2$	MSE masseurs the average of the squared deviation between the actual and the estimated value.	MSE is a positive value. It is desired to be minimum where the deviation between the predicted and the actual data is low. When MSE is zero, the model is ideal.
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i)^2}$	RMES denotes evaluating the results in the short-term testing.	RMES is a positive value between zero and infinity. It is desired that the value is near zero.
APE	$\left\lvert\frac{y_i-x_i}{y_i}\right\rvert\times 100$	APE define the difference between the actual and forecasted values.	It is preferred to have lower APE value as zero indicate no error.

y_i is the forecasted value, x_i is the actual value, n is the number of instances and $\overline{x_i}$ is the mean of the actual data (Ağbulut et al., 2021).

4. IMPLEMENTATION AND RESULTS

The data were collected and exported to WEKA as arff file to train and test the time series forecasting algorithms. After reaching acceptable models by evaluating the results using the statistical metrics the forecasted results were collected and compared between the two different methods was made to distinguish which of them had better performance.

4.1. Electricity consumption data

The collected data from Turkstat (TurkStat - Data Portal, 2022) consist of the total electricity consumption and the electricity consumption rate for each sector in the period 1970-2020. The sectoral electricity consumption in GWh is shown in Table 4.1.

Table 4.1 Total and sectoral electricity consumption (TurkStat - Data Portal, 2022)

Year	Household	Commercial	Government	Industrial	Illumination	Other	Total
1970	1162.0	350.8	299.6	4691.7	190.0	613.9	7308.0
1971	1351.1	381.3	339.8	5346.4	198.9	671.4	8289.0
1972	1533.8	447.8	362.0	6192.6	209.6	781.2	9527.0
1973	1558.4	452.8	368.6	7086.7	221.1	842.4	10530.0
1974	1726.6	579.3	431.6	7576.5	227.2	817.8	11359.0
1975	2361.1	661.1	499.2	8742.8	256.3	971.4	13492.0
1976	2813.8	755.7	562.8	10499.6	257.3	1189.8	16079.0
1977	3180.5	898.5	557.0	11985.3	251.6	1096.1	17969.0
1978	3578.5	927.8	605.9	12401.8	284.0	1136.0	18934.0
1979	3946.2	1119.1	628.3	12545.5	294.5	1099.4	19633.0
1980	4385.6	1142.3	611.9	13013.9	285.6	958.7	20398.0
1981	4604.3	1255.7	638.9	14209.4	308.4	1013.4	22030.0
1982	4929.7	1368.0	589.7	15190.0	306.6	1202.9	23587.0
1983	5137.7	1394.5	685.0	15584.2	293.6	1370.0	24465.0
1984	5471.7	1575.2	773.8	18018.0	331.6	1464.7	27635.0
1985	5644.7	1634.0	891.3	19607.9	415.9	1515.2	29709.0
1986	6119.9	1674.9	1030.7	20872.1	676.4	1836.0	32210.0
1987	6935.7	1761.5	1174.3	23889.7	770.6	2165.1	36697.0
1988	7944.4	1986.1	1271.1	25263.2	834.2	2423.0	39722.0
1989	8451.5	2285.4	1293.6	27596.8	905.5	2587.2	43120.0
1990	9176.7	2575.1	1451.4	29215.7	1217.3	3183.8	46820.0
1991	10842.3	3055.5	1872.8	28534.9	1429.2	3548.4	49283.0
1992	11498.8	3293.1	1997.4	31527.2	1835.5	3832.9	53985.0
1993	12558.2	3613.5	2251.0	34239.0	2251.0	4324.3	59237.0
1994	13446.8	3684.1	3315.7	34139.0	2517.4	4298.1	61401.0
1995	14489.7	4178.4	3032.7	38010.2	3100.1	4582.8	67394.0
1996	16388.7	5710.1	2966.3	40638.0	3114.6	5339.3	74157.0
1997	18506.0	6878.3	3766.7	43480.9	3275.4	5977.6	81885.0
1998	19996.7	7718.0	4297.5	46132.8	3683.6	5876.2	87705.0
1999	22618.1	8208.2	3739.3	46513.0	4195.3	5928.1	91202.0
2000	23885.9	9338.1	4128.4	48853.1	4521.6	7568.8	98296.0
2001	23588.0	9901.1	4368.2	46981.9	4853.5	7377.3	97070.0
2002	23575.1	10912.5	4529.7	50444.5	5147.4	8338.8	102948.0
2003	25147.4	12853.1	4582.4	55100.6	5029.5	9053.0	111766.0
2004	27620.4	15627.3	4482.3	59601.9	4482.3	9327.9	121142.0
2005	30872.3	18497.3	4689.5	62265.7	4168.4	9769.7	130263.0
2006	34480.1	20316.1	6009.0	67958.7	4006.0	10301.1	143071.0
2007	36456.7	23115.1	6981.1	73844.3	4033.5	10704.3	155135.0
2008	39515.3	23968.3	7287.7	74820.0	4048.7	12308.0	161948.0
2009	39223.5	24946.1	7060.2	70445.4	3922.4	11296.4	156894.0
2010	41464.3	27700.2	7054.1	79315.5	3785.1	12731.8	172051.0
2011	44291.8	30520.4	7257.9	88025.3	3908.1	12096.5	186100.0
2012	45417.1	31772.4	8771.5	92393.5	3898.5	12670.0	194923.0
2013	44956.2	37430.5	8119.8	93279.2	3762.9	10496.4	198045.0
2014	46244.6	39816.0	8087.6	97881.0	3940.1	11405.6	207375.0
2015	47808.6	41506.6	8040.5	103440.5	4128.9	12386.8	217312.0
2016	51327.2	43466.3	9016.9	108434.5	4161.7	14797.0	231203.7
2017	54286.9	49306.5	10209.9	116542.6	4482.4	14194.3	249022.6
2018	54487.0	52679.3	11878.7	117753.8	4648.2	16785.1	258232.0
2019	56194.1	49597.4	13660.3	115675.4	5074.9	17071.1	257273.1
2020	60693.5	45476.4	12815.5	120027.3	5237.5	18451.9	262702.1

Figure 4.1 shows the sectoral and total electricity consumption through the period 1970-2020.

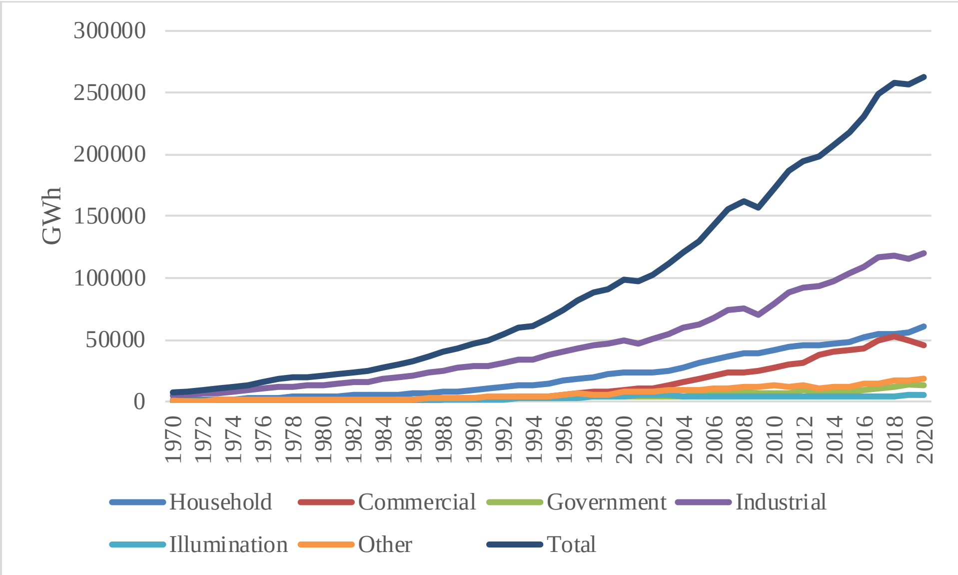

Figure 4.1 Total Electricity Consumption

The electricity consumption rates for each sector of the total electricity consumption are shown in Table 4.2.

Table 4.2 Sectoral share of total electricity consumption (TurkStat - Data Portal, 2022)

Year	Household	Commercial	Government	Industrial	Illumination	Other
1970	15.9	4.8	4.1	64.2	2.6	8.4
1971	16.3	4.6	4.1	64.5	2.4	8.1
1972	16.1	4.7	3.8	65.0	2.2	8.2
1973	14.8	4.3	3.5	67.3	2.1	8.0
1974	15.2	5.1	3.8	66.7	2.0	7.2
1975	17.5	4.9	3.7	64.8	1.9	7.2
1976	17.5	4.7	3.5	65.3	1.6	7.4
1977	17.7	5.0	3.1	66.7	1.4	6.1
1978	18.9	4.9	3.2	65.5	1.5	6.0
1979	20.1	5.7	3.2	63.9	1.5	5.6
1980	21.5	5.6	3.0	63.8	1.4	4.7
1981	20.9	5.7	2.9	64.5	1.4	4.6
1982	20.9	5.8	2.5	64.4	1.3	5.1
1983	21.0	5.7	2.8	63.7	1.2	5.6
1984	19.8	5.7	2.8	65.2	1.2	5.3
1985	19.0	5.5	3.0	66.0	1.4	5.1
1986	19.0	5.2	3.2	64.8	2.1	5.7
1987	18.9	4.8	3.2	65.1	2.1	5.9
1988	20.0	5.0	3.2	63.6	2.1	6.1
1989	19.6	5.3	3.0	64.0	2.1	6.0
1990	19.6	5.5	3.1	62.4	2.6	6.8
1991	22.0	6.2	3.8	57.9	2.9	7.2
1992	21.3	6.1	3.7	58.4	3.4	7.1
1993	21.2	6.1	3.8	57.8	3.8	7.3
1994	21.9	6.0	5.4	55.6	4.1	7.0
1995	21.5	6.2	4.5	56.4	4.6	6.8
1996	22.1	7.7	4.0	54.8	4.2	7.2
1997	22.6	8.4	4.6	53.1	4.0	7.3
1998	22.8	8.8	4.9	52.6	4.2	6.7
1999	24.8	9.0	4.1	51.0	4.6	6.5
2000	24.3	9.5	4.2	49.7	4.6	7.7
2001	24.3	10.2	4.5	48.4	5.0	7.6
2002	22.9	10.6	4.4	49.0	5.0	8.1
2003	22.5	11.5	4.1	49.3	4.5	8.1
2004	22.8	12.9	3.7	49.2	3.7	7.7
2005	23.7	14.2	3.6	47.8	3.2	7.5
2006	24.1	14.2	4.2	47.5	2.8	7.2
2007	23.5	14.9	4.5	47.6	2.6	6.9
2008	24.4	14.8	4.5	46.2	2.5	7.6
2009	25.0	15.9	4.5	44.9	2.5	7.2
2010	24.1	16.1	4.1	46.1	2.2	7.4
2011	23.8	16.4	3.9	47.3	2.1	6.5
2012	23.3	16.3	4.5	47.4	2.0	6.5
2013	22.7	18.9	4.1	47.1	1.9	5.3
2014	22.3	19.2	3.9	47.2	1.9	5.5
2015	22.0	19.1	3.7	47.6	1.9	5.7
2016	22.2	18.8	3.9	46.9	1.8	6.4
2017	21.8	19.8	4.1	46.8	1.8	5.7
2018	21.1	20.4	4.6	45.6	1.8	6.5
2019	21.8	19.3	5.3	45.0	2.0	6.6
2020	23.1	17.3	4.9	45.7	2.0	7.0

Figure 4.2 present the changes in the rates for electricity consumption during the period 1970-2020. Through the years the household and commercial sectors had a linear increase in the share of the total electricity consumption. However, the industrial sector witnessed a linear decrease as the rest of the sectors stayed around the same electricity consumption ratio.

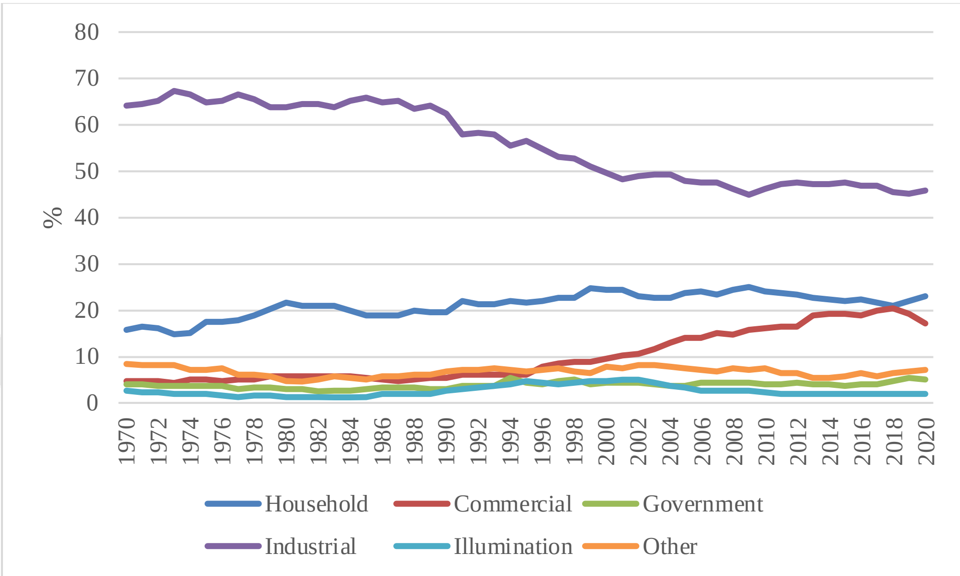

Figure 4.2 Sectoral shares of total electricity consumption

4.2. Creating the data files

The collected data from Turkstat (TurkStat - Data Portal, 2022) contains the total electricity consumption in GWh and the percentage of each sector every year from 1970 until 2020. Using the percentage, the amount of electricity consumption for each sector in GWh was calculated.

The obtained data was presented in an Excel file so to use the data in WEKA another type of file had to be created.

There is multiple file format that can be used in WEKA. In this study, the attribute relation file format (arff) was used where the name of the relation between the attributes is defined as well as a list of the attributes and their type (Date, numeric or nominal) after that the data information is presented.

Two separate arff files were created to be used in WEKA to forecast the electricity consumption.

The first one contains seven attributes, the year, and six attributes each one is the electricity consumption for a specified sector in GWh where the year attribute was identified as a date attribute and the consumption was identified as a numeric attribute. The second file contained two attributes, the year, and the total electricity consumption for that year where the year attribute was identified as a date attribute and the electricity consumption was identified as a numeric attribute.

The Data file arff file structure:

@relation Data

@attribute year date 'yyyy'

@attribute household numeric

@attribute commercial numeric

@attribute government numeric

@attribute industrial numeric

@attribute illumination numeric

@attribute other numeric

@data

1970 1161.972 350.784 299.628 4691.736 190.008 613.872

1971 1351.107 381.294 339.849 5346.405 198.936 671.409

1972 1533.847 447.769 362.026 6192.55 209.594 781.214

The total arff file structure:

@relation Total

@attribute year date 'yyyy'

@attribute total numeric

@data

1970 7308

1971 8289

4.3. Machine Learning parameters

Various parameters were tested to reach the highest performance and accuracy for the forecasting model created by the machine learning algorithms used in this study.

Each attribute was tested individually to ascertain that no over-fitting or under-fitting occurred in addition to obtaining the parameters that meet the desired precision.

4.3.1. SMOReg parameters

There are multiple parameters used to create the model using SMOReg some of them are the same for all the attributes while other parameters change according to each attribute as shown in Table 4.3.

Table 4.3 SMOReg parameters

Parameters	Commercial	Government	Household	Illumination	Industrial	other	Total
Complexity	22						
Filter type	Normalize training data						
Kernel Function	PolyKernel						
Regression Optimizer	RegSMOImproved						
Training data%	70%	80%	80%	80%	80%	80%	80%
Testing data%	30%	20%	20%	20%	20%	20%	20%
Lag length	4	4	4	4	4	4	5

4.3.2. MLPRegressor parameters

Creating the forecasting model using MLPRegressor requires different parameters. These parameters differ from one attribute to another in order to get the best forecast model for each attribute. The whole data were used to train the algorithm as MLPRegressor has a high dependence on the number of neurons between the input and the hidden layer. The attributes were normalized, and the parameters are shown in Table 4.4.

Table 4.4 MLPRegressor parameters

Parameters	Commercial	Government	Household	Illumination	Industrial	Other	Total
Momentum	0.2						
Hidden layers	9	9	9	10	10	10	9
Learning rate	0.3	0.3	0.4	0.4	0.4	0.4	0.3
Lag length	10	11	9	12	20	11	13

The lag length establishes how many previous instances affect over the forecasted value.

The learning rate is the rate at which the weight of the inputs updates. And the momentum is the momentum applied to the weight updates.

The value of the hidden layers determines how many nodes are inside the hidden layer in between the input and the output layers as shown in Figure 4.3.

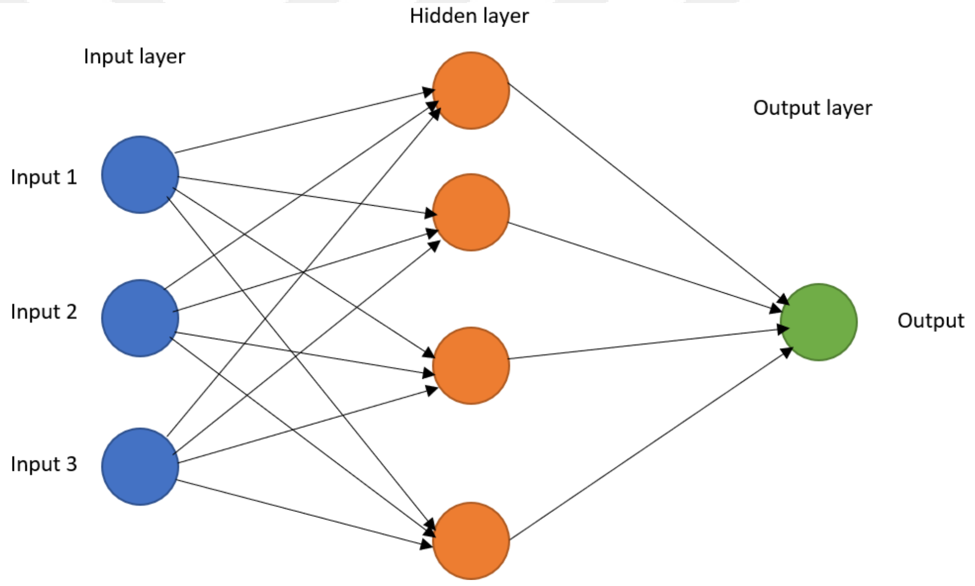

Figure 4.3 MLP layers

4.4. Obtained formulas

These formulas are generated by the machine learning algorithms whether SMOReg or MLPRegressor and used to forecast the next value of the electricity consumption in each sector.

4.4.1. SMOReg formulas

The generated formulas by using SMOReg are different for each attribute depending on the parameters and the data values that trained the algorithm illustrated in Equation (13).

$$
\begin{aligned}
a \times year\ remaped + \sum_{n=1}^{n} b_n \times lag_n \ + \ c \times year\ remaped^2 \\
+ \ d \times year\ remaped^3 \\
+ \sum_{n=1}^{n} e_n \times year\ remaped \times lag_n + f
\end{aligned}
\tag{13}
$$

Where *a*, *b*, *c*, *d*, *e*, and *f* are constants generated by training SMOReg algorithm. The appendix shows the SMOReg generated formula for each attribute.

4.4.2. MLPRegressor formulas

MLPRegressor produces a formula after training the algorithm according to the stated parameters.

The formulas generated by MLPRegressor are complicated where for each hidden node there is a formula and the higher the number of the nodes as well as the number of interconnections between the nodes and the other layers are the more complicated the formula is.

As the number of nodes for the hidden layer is 9 and higher. the MLPregressor formula can be represented by the visualization of the layers.

Figure 4.4 represent the relation between the layers that forecast the commercial sector electricity consumption where there are nine nodes, and the past ten instances affect the forecasted value.

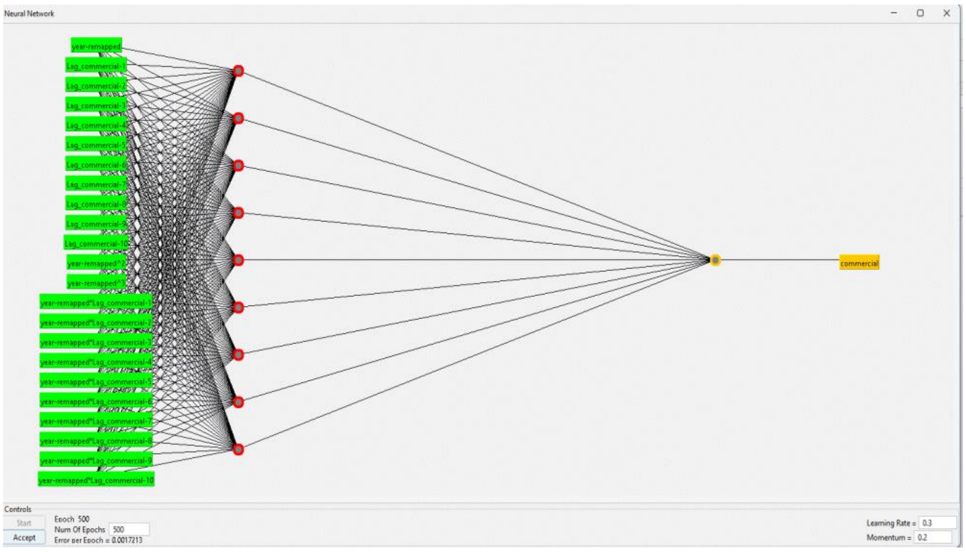

Figure 4.4 MLP Forecasting commercial electricity consumption layers

Figure 4.5 represents the relation between the layers that forecast the government sector electricity consumption where the hidden layer contains nine nodes, and the forecasted value is affected by the past eleven instances.

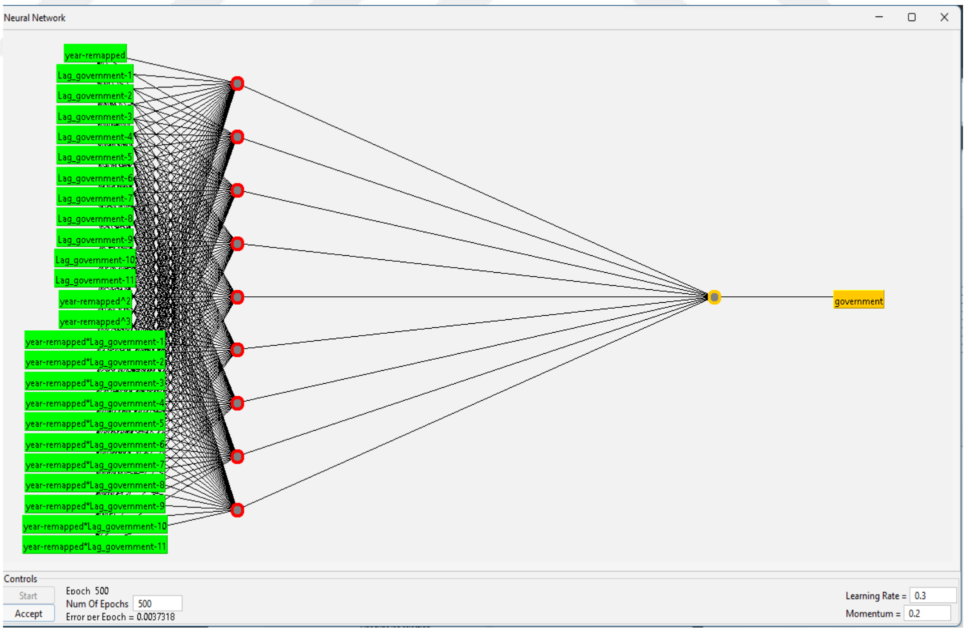

Figure 4.5 MLP Forecasting government electricity consumption layers

Figure 4.6 represent the relation between the layers that forecast the household sector electricity consumption as the forecasted value is influenced by the past nine years and the hidden layer contains 9 nodes.

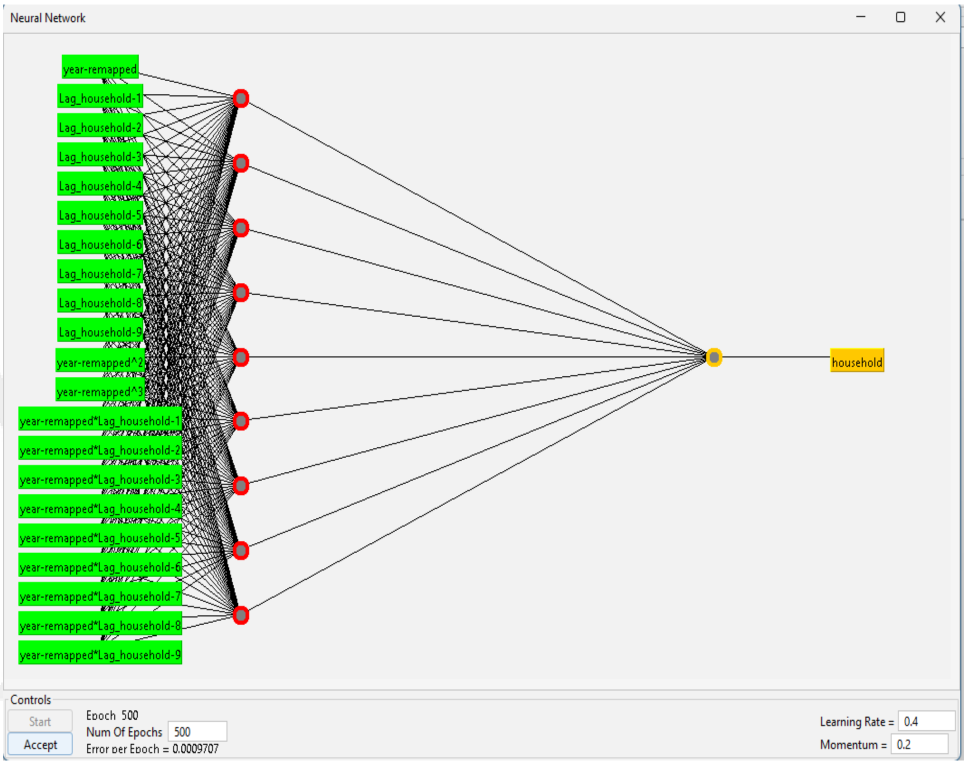

Figure 4.6 MLP Forecasting household electricity consumption layers

Figure 4.7 represent the relation between the layers that forecast the illumination sector electricity consumption where the hidden layer has ten nodes and the past twelve years have an effect on the forecasted value.

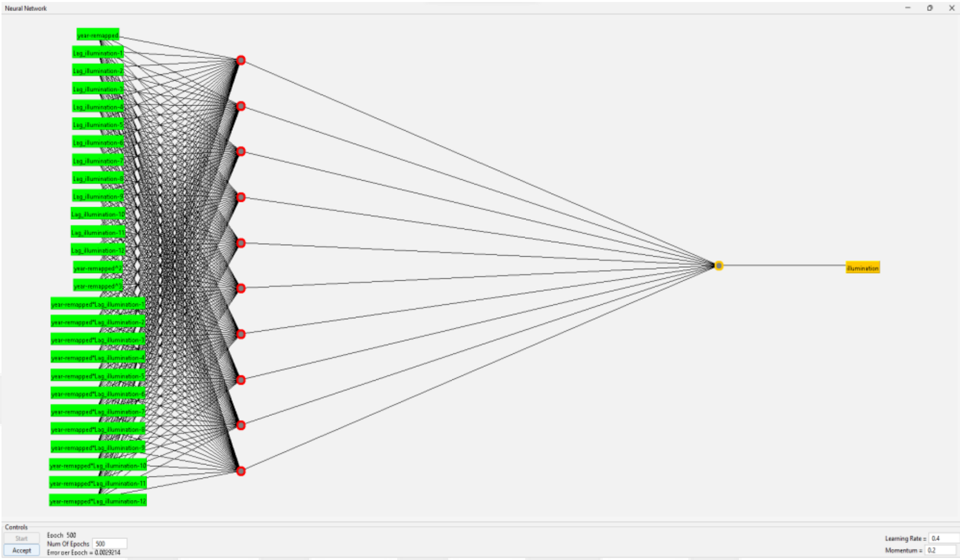

Figure 4.7 MLP Forecasting illumination electricity consumption layers

Figure 4.8 represent the relation between the layers that forecast the industry sector electricity consumption as the hidden layer consists of ten nodes where the past twenty years influence the forecasted value.

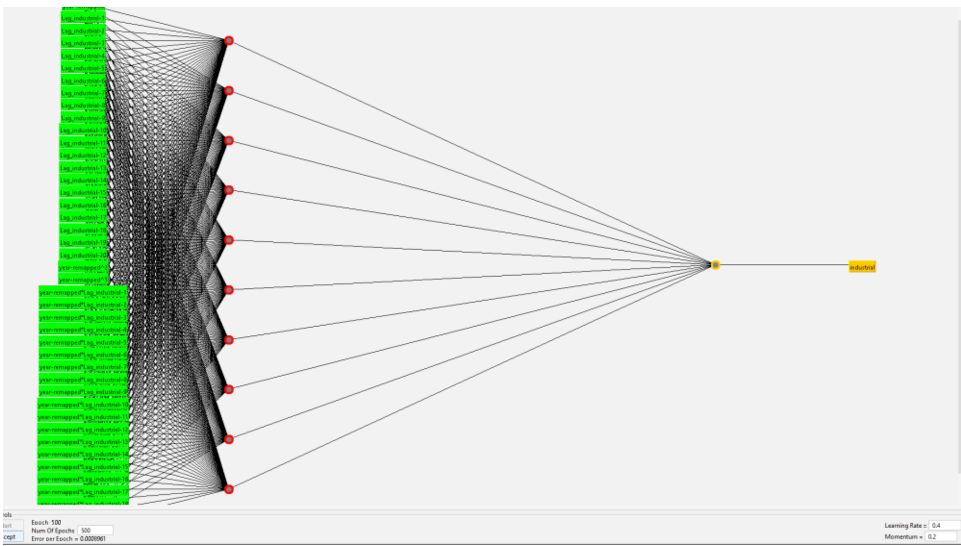

Figure 4.8 MLP Forecasting industrial electricity consumption layers

Figure 4.9 represent the relation between the layers that forecast the other sectors' electricity consumption where the past eleven years affect the result of the forecasted value and the hidden layer has ten nodes.

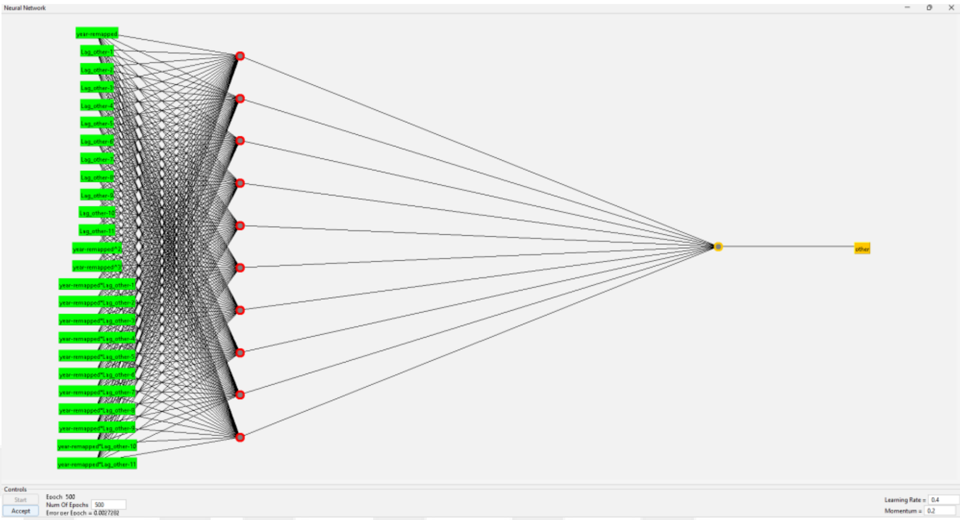

Figure 4.9 MLP Forecasting other sectors' electricity consumption layers

Figure 4.10 represent the relation between the layers that forecast the total electricity consumption. The forecasted total electricity is influenced by the past thirteen years and the hidden layer has nine nodes.

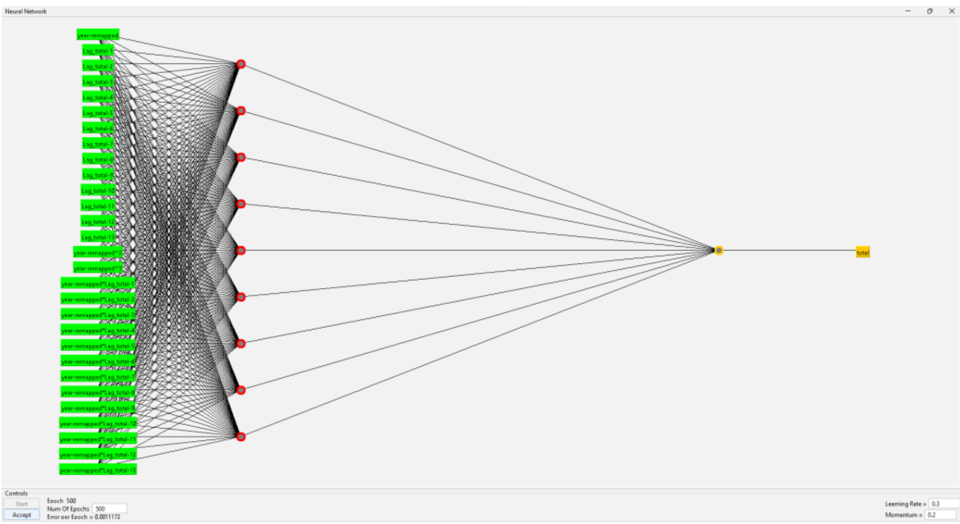

Figure 4.10 MLP Forecasting total electricity consumption layers

4.5. Forecasted results

Using SMOReg and MLPRegressor the electricity consumption for the years 2021 until the year 2050 as the given data had the actual values of electricity consumption from the year 1970 until 2020.

Each sector has been forecasted individually and the summation of the forecasted sectors is presented as well as the forecasted total electricity consumption.

4.5.1. SMOReg results

The forecasted electricity consumption for the sectors in GWh is shown in Table 4.5 where the total consumption here is the summation of the forecasted electricity consumption for the sectors in that year.

Table 4.5 Forecasted sectoral electricity consumption using SMOReg

year	Household	Commercial	Government	Industrial	Illumination	Other	total
2021	63940.80	58785.13	10391.18	126226.80	4213.48	20853.02	284410.41
2022	66688.33	68944.63	10133.31	129818.85	3406.64	20719.02	299710.78
2023	68457.67	77880.88	12319.12	133059.75	3494.64	21606.85	316818.91
2024	69741.63	86866.29	14827.17	138521.75	4123.74	22071.21	336151.79
2025	71965.81	94516.39	15408.16	144211.96	4560.94	22747.13	353410.38
2026	75140.14	101698.65	14082.56	149007.93	4481.35	23715.76	368126.40
2027	78750.08	108839.80	13008.84	154042.13	4167.45	24435.74	383244.04
2028	81972.57	116048.58	13986.02	159994.15	4076.64	25304.55	401382.53
2029	84321.53	123573.03	16431.73	165925.98	4327.90	26076.11	420656.28
2030	86210.83	131492.36	18040.41	171601.18	4637.58	26870.58	438852.95
2031	88486.75	139835.29	17452.46	177605.28	4702.06	27714.46	455796.31
2032	91664.09	148625.44	15918.14	184059.87	4551.68	28536.46	473355.67
2033	95484.52	157871.71	15956.58	190514.77	4479.01	29408.28	493714.87
2034	99065.32	167555.22	18349.85	196972.52	4656.71	30260.56	516860.17
2035	101793.57	177677.13	20977.95	203746.83	4923.84	31132.60	540251.92
2036	103907.90	188240.67	21141.31	210789.98	5004.18	32015.16	561099.20
2037	106336.72	199264.17	18973.19	217900.03	4878.55	32909.12	580261.78
2038	109774.48	210741.45	17744.27	225124.35	4819.72	33814.35	602018.61
2039	114001.35	222686.12	20111.24	232640.94	5022.00	34728.02	629189.67
2040	117984.10	235113.09	24383.93	240355.61	5309.64	35654.10	658800.47
2041	120892.81	248050.24	25831.54	248203.92	5356.95	36596.06	684931.52
2042	122958.06	261493.93	22486.61	256213.40	5166.89	37539.89	705858.77
2043	125442.20	275460.47	18696.59	264500.08	5124.38	38500.45	727724.18
2044	129384.12	289966.06	20768.16	272960.97	5452.54	39467.54	757999.39
2045	134431.43	305040.60	28443.27	281615.27	5807.00	40454.78	795792.34
2046	138914.39	320678.37	33051.35	290445.57	5709.32	41441.53	830240.52
2047	141562.16	336895.96	27420.25	299552.32	5334.85	42445.59	853211.12
2048	143023.12	353710.18	17254.39	308829.22	5410.20	43456.40	871683.51
2049	145572.67	371153.40	17558.83	318354.07	6094.80	44487.22	903220.99
2050	150730.02	389218.48	33325.96	328050.20	6475.55	45517.06	953317.27

According to the summed forecasted electricity consumption for the sectors the percentage of the consumption for each sector was calculated and presented in Table 4.6.

Table 4.6 Forecasted sectoral electricity consumption share using SMOReg

year	Household	Commercial	Government	Industrial	Illumination	Other
2021	22.48	20.67	3.65	44.38	1.48	7.33
2022	22.25	23.00	3.38	43.31	1.14	6.91
2023	21.61	24.58	3.89	42.00	1.10	6.82
2024	20.75	25.84	4.41	41.21	1.23	6.57
2025	20.36	26.74	4.36	40.81	1.29	6.44
2026	20.41	27.63	3.83	40.48	1.22	6.44
2027	20.55	28.40	3.39	40.19	1.09	6.38
2028	20.42	28.91	3.48	39.86	1.02	6.30
2029	20.05	29.38	3.91	39.44	1.03	6.20
2030	19.64	29.96	4.11	39.10	1.06	6.12
2031	19.41	30.68	3.83	38.97	1.03	6.08
2032	19.36	31.40	3.36	38.88	0.96	6.03
2033	19.34	31.98	3.23	38.59	0.91	5.96
2034	19.17	32.42	3.55	38.11	0.90	5.85
2035	18.84	32.89	3.88	37.71	0.91	5.76
2036	18.52	33.55	3.77	37.57	0.89	5.71
2037	18.33	34.34	3.27	37.55	0.84	5.67
2038	18.23	35.01	2.95	37.39	0.80	5.62
2039	18.12	35.39	3.20	36.97	0.80	5.52
2040	17.91	35.69	3.70	36.48	0.81	5.41
2041	17.65	36.22	3.77	36.24	0.78	5.34
2042	17.42	37.05	3.19	36.30	0.73	5.32
2043	17.24	37.85	2.57	36.35	0.70	5.29
2044	17.07	38.25	2.74	36.01	0.72	5.21
2045	16.89	38.33	3.57	35.39	0.73	5.08
2046	16.73	38.62	3.98	34.98	0.69	4.99
2047	16.59	39.49	3.21	35.11	0.63	4.97
2048	16.41	40.58	1.98	35.43	0.62	4.99
2049	16.12	41.09	1.94	35.25	0.67	4.93
2050	15.81	40.83	3.50	34.41	0.68	4.77

Figure 4.11 shows the fluctuation of the forecasted percentage for each sectors of total electricity consumption

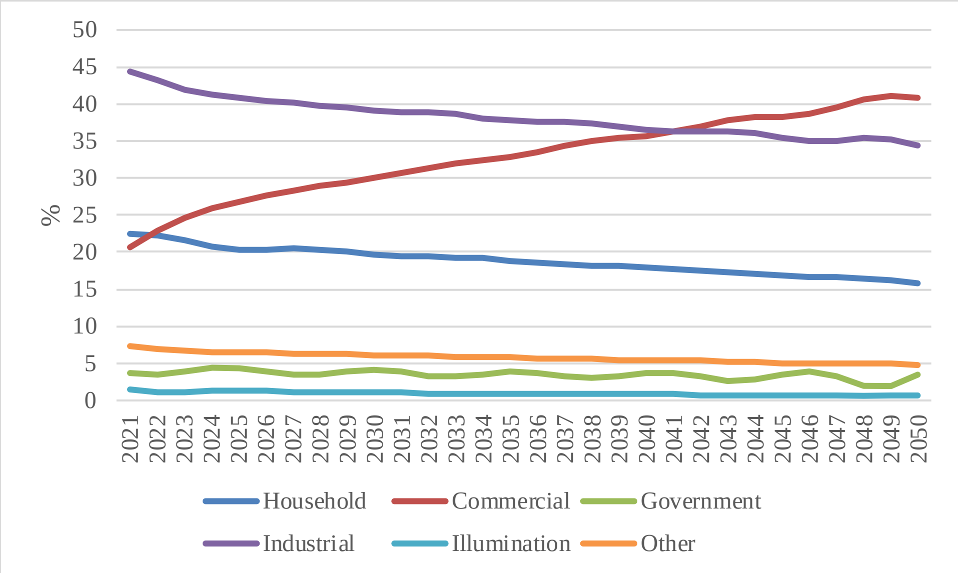

Figure 4.11 Forecasted sectoral shares of total electricity consumption using SMOReg

The forecasted energy consumption for the commercial sector in 2021 was 58758.13 GWh and reached 389218.47 GWh in 2050 where it accumulates 20.66% and 40.82% respectively of the total electricity consumption as illustrated in Figure 4.12.

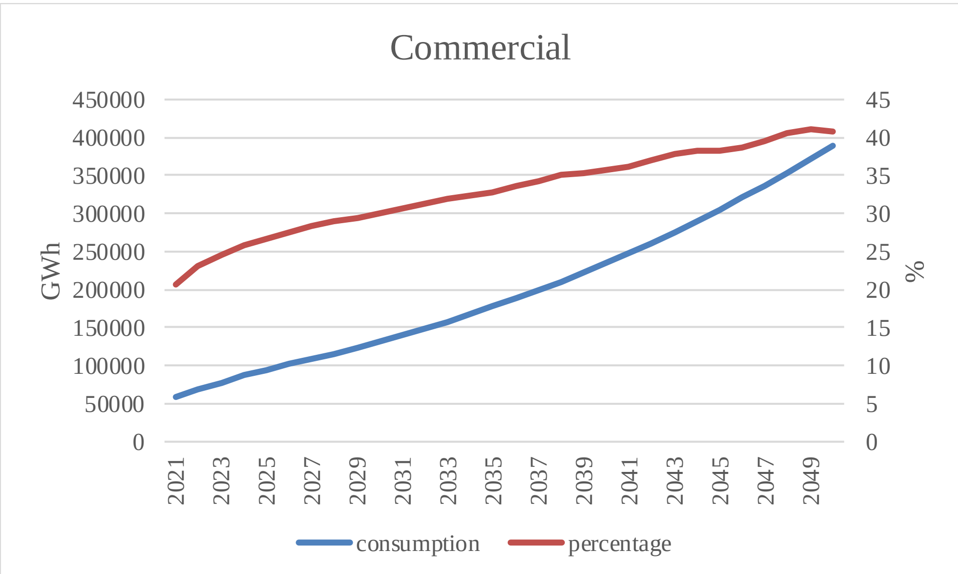

Figure 4.12 Forecasted commercial electricity consumption using SMOReg

Figure 4.13 demonstrate the changes in the electricity consumption in the government sector where it was 3.65% of the total electricity consumption with 10391.14 GWh electricity consumption in 2021 and reaching 3.49% and 33325.95 GWh in 2050.

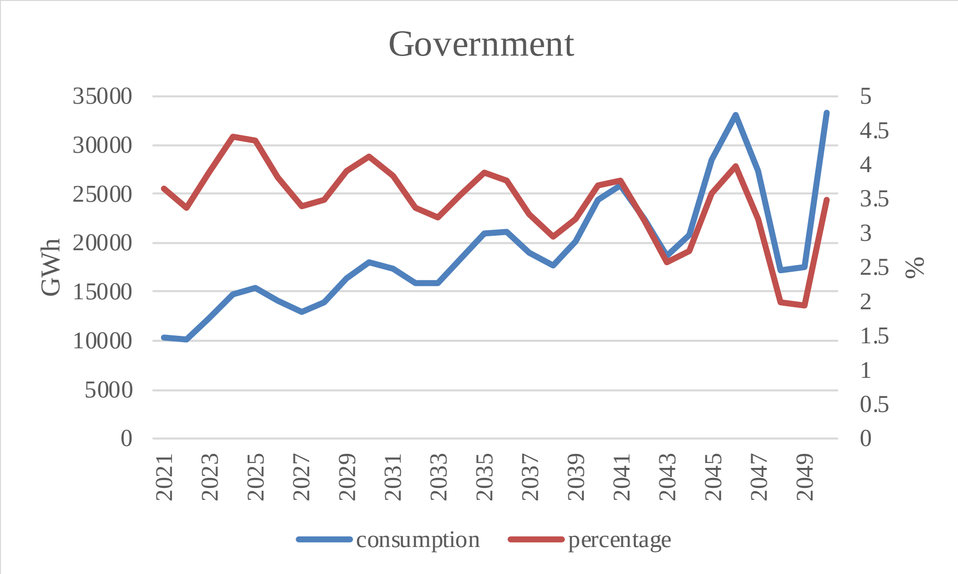

Figure 4.13 Forecasted government electricity consumption using SMOReg

The household sector registered a rising in the electricity consumption from 63940.8 GWh in 2021 to 150730.02 GWh in 2050 where the share of the household sector of the total electricity consumption dropped from 22.48% to 15.81% as presented in Figure 4.14.

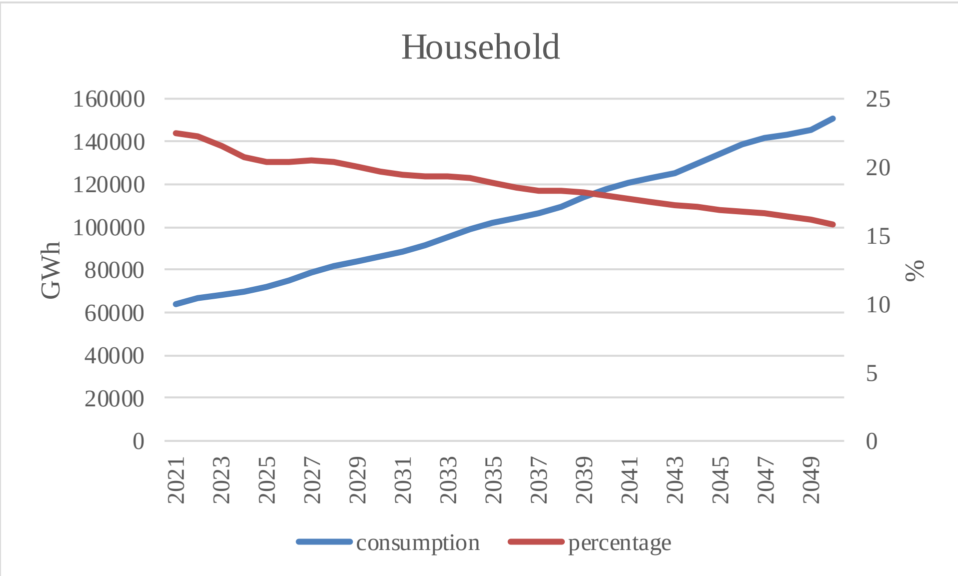

Figure 4.14 Forecasted household electricity consumption using SMOReg

The illumination sector registered a rising in the electricity consumption which went from 4213.47 GWh to 6475.55 GWh where the share of the total consumption went down from 1.48% to 0.67% in the period 2021-2050 as shown in Figure 4.15.

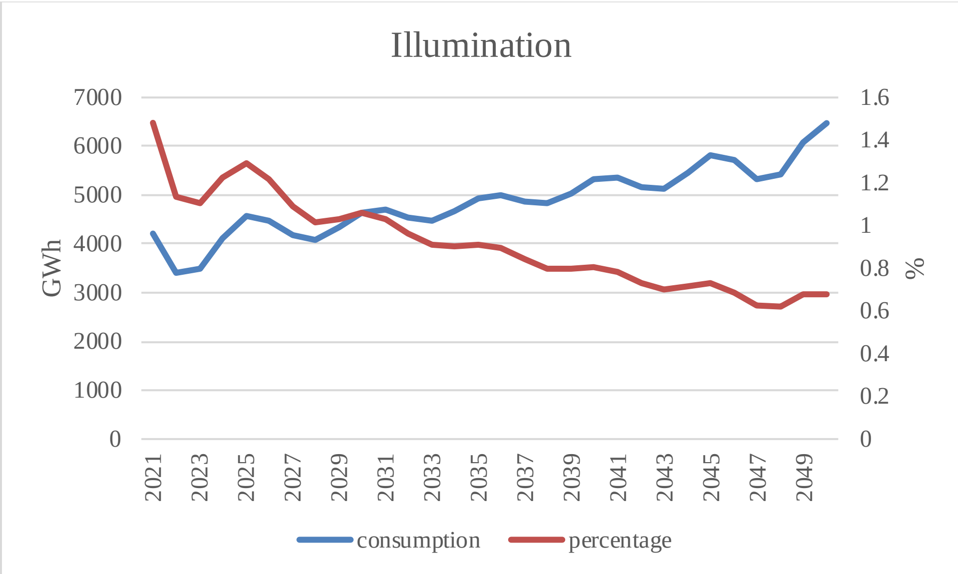

Figure 4.15 Forecasted illumination electricity consumption using SMOReg

The industrial sector came with the highest electricity consumption through the forecasted period as in 2021 the industrial electricity consumption was 126226.8 GWh with 44.38% share of the total electric consumption in that year, in 2050 the electricity consumption for the industrial sector went up to 328050.19 GWh with 34.4% of the total electricity consumption as shown in Figure 4.16.

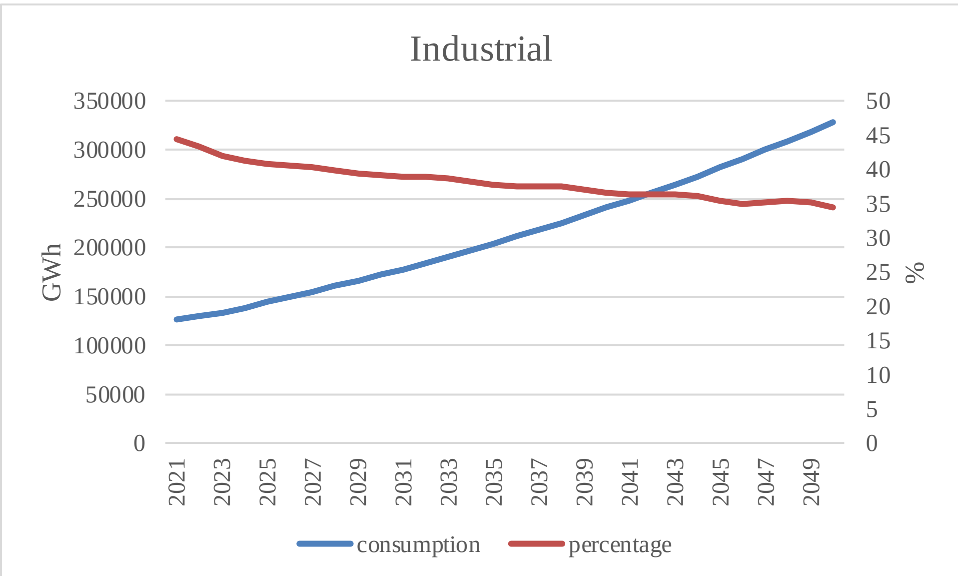

Figure 4.16 Forecasted industrial electricity consumption using SMOReg

The other sectors combined had 20179 GWh in 2021 and reached 45517 GWh in 2050 where the percentage was 7.33% and 4.77% respectively as demonstrated in Figure 4.17.

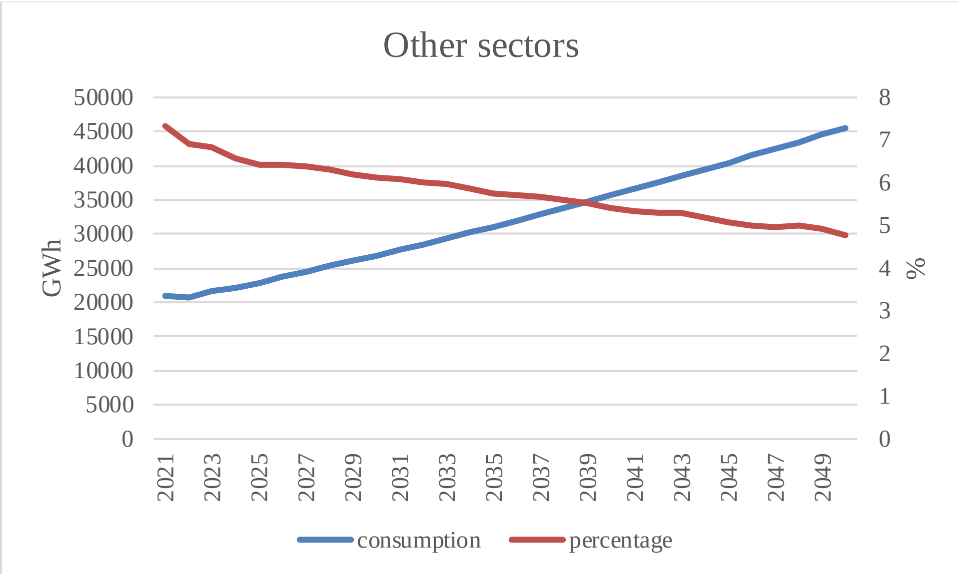

Figure 4.17 Forecasted other sectors' electricity consumption using SMOReg

The forecasted total electricity consumption is shown in Table 4.7.

Table 4.7 Forecasted total electricity consumption using SMOreg

Year	Total	Year	Total	Year	Total
2021	292147.2	2031	426722.8	2041	596747.5
2022	302047.9	2032	442316.8	2042	615489.5
2023	312563.8	2033	458226.6	2043	634609.1
2024	325959.9	2034	474256.5	2044	654051.5
2025	341838.8	2035	490843.4	2045	673928
2026	353485.6	2036	507651.4	2046	693992.5
2027	367441.4	2037	524857.2	2047	714440.8
2028	382056.9	2038	542275.4	2048	735227.3
2029	396972.3	2039	560099.5	2049	756456.3
2030	411306.2	2040	578211.3	2050	777854.4

4.5.2. MLPRegressor results

Table 4.8 shows the forecasted values of the electricity consumption in GWh by sectors and the total value in this table represents the summation of the forecasted values for the electricity consumption for the sectors in that year.

Table 4.8 Forecasted sectoral electricity consumption using MLPRegressor

Year	Commercial	Government	Household	Illumination	Industrial	Other	Total
2021	46072.69	11724.03	61709.15	5247.72	132069.44	14095.55	270918.58
2022	43359.63	12206.96	62908.89	5376.48	130351.71	15187.31	269390.96
2023	41256.41	13386.76	64953.32	5423.52	126059.73	15726.23	266805.96
2024	39736.44	13527.61	66483.23	5407.79	131465.59	18474.89	275095.54
2025	39520.48	13886.86	68736.51	5303.61	135577.14	16676.22	279700.81
2026	38116.22	14030.61	71516.34	5188.02	134042.80	19791.23	282685.22
2027	38470.72	14520.16	72858.57	4994.52	133800.24	21086.07	285730.28
2028	42037.07	14935.30	74410.02	4666.87	133369.24	22228.44	291646.95
2029	44223.76	15315.03	77076.74	4443.69	132088.80	20064.04	293212.05
2030	42766.01	15748.09	78783.95	4372.39	135165.52	21384.09	298220.04
2031	44151.38	15779.99	80184.18	4431.06	137473.55	21433.13	303453.28
2032	45901.73	15561.82	81721.86	4575.39	131845.82	21704.32	301310.94
2033	45870.14	15531.41	82937.56	4854.70	129676.69	20452.35	299322.84
2034	43577.50	15601.17	84126.76	5336.42	135886.88	22083.53	306612.25
2035	43263.36	15540.61	85338.14	5937.87	138149.42	22650.37	310879.77
2036	42603.63	15447.86	86209.45	6481.76	133333.24	23077.11	307153.05
2037	40239.42	15299.79	86986.67	6779.03	132462.59	22923.91	304691.41
2038	39632.89	15195.96	87792.61	6868.29	137598.48	23486.65	310574.88
2039	41125.93	15098.44	88436.30	6862.99	140670.69	23569.42	315763.78
2040	40629.70	15061.80	88978.14	6818.06	140719.62	23597.71	315805.03
2041	39869.77	15150.97	89462.11	6767.82	142385.00	23582.07	317217.73
2042	41044.18	15259.77	89865.90	6723.28	143534.82	23763.37	320191.31
2043	43715.63	15317.36	90215.94	6680.72	148236.59	23816.49	327982.73
2044	43421.40	15400.47	90522.49	6598.83	156246.57	23869.09	336058.85
2045	43014.93	15527.76	90780.69	6411.54	158038.52	23922.45	337695.89
2046	44457.25	15661.98	91002.38	6160.67	156495.18	23969.64	337747.10
2047	44419.76	15785.98	91196.86	6019.30	162574.06	23995.53	343991.50
2048	42431.36	15888.50	91365.18	6118.37	169561.52	24013.80	349378.74
2049	42233.30	15971.95	91512.43	6440.74	170481.30	24030.26	350669.99
2050	42893.02	16022.58	91641.70	6805.97	170570.70	24041.59	351975.56

The share that each sector had of the total electricity consumption is presented in Table 4.9.

Table 4.9 Forecasted sectoral electricity consumption share using MLPRegressor

Year	Commercial	Government	Household	Illumination	Industrial	Other
2021	17.01	4.33	22.78	1.94	48.75	5.20
2022	16.10	4.53	23.35	2.00	48.39	5.64
2023	15.46	5.02	24.34	2.03	47.25	5.89
2024	14.44	4.92	24.17	1.97	47.79	6.72
2025	14.13	4.96	24.58	1.90	48.47	5.96
2026	13.48	4.96	25.30	1.84	47.42	7.00
2027	13.46	5.08	25.50	1.75	46.83	7.38
2028	14.41	5.12	25.51	1.60	45.73	7.62
2029	15.08	5.22	26.29	1.52	45.05	6.84
2030	14.34	5.28	26.42	1.47	45.32	7.17
2031	14.55	5.20	26.42	1.46	45.30	7.06
2032	15.23	5.16	27.12	1.52	43.76	7.20
2033	15.32	5.19	27.71	1.62	43.32	6.83
2034	14.21	5.09	27.44	1.74	44.32	7.20
2035	13.92	5.00	27.45	1.91	44.44	7.29
2036	13.87	5.03	28.07	2.11	43.41	7.51
2037	13.21	5.02	28.55	2.22	43.47	7.52
2038	12.76	4.89	28.27	2.21	44.30	7.56
2039	13.02	4.78	28.01	2.17	44.55	7.46
2040	12.87	4.77	28.18	2.16	44.56	7.47
2041	12.57	4.78	28.20	2.13	44.89	7.43
2042	12.82	4.77	28.07	2.10	44.83	7.42
2043	13.33	4.67	27.51	2.04	45.20	7.26
2044	12.92	4.58	26.94	1.96	46.49	7.10
2045	12.74	4.60	26.88	1.90	46.80	7.08
2046	13.16	4.64	26.94	1.82	46.34	7.10
2047	12.91	4.59	26.51	1.75	47.26	6.98
2048	12.14	4.55	26.15	1.75	48.53	6.87
2049	12.04	4.55	26.10	1.84	48.62	6.85
2050	12.19	4.55	26.04	1.93	48.46	6.83

Figure 4.18 represent the sector's share in the total electricity consumption between 2021 and 2050.

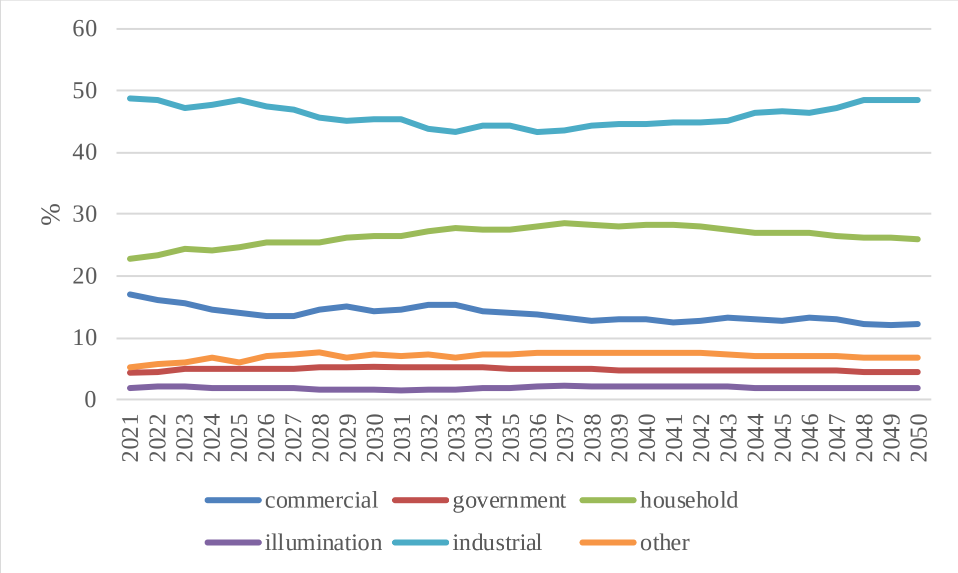

Figure 4.18 Forecasted sectoral rates of total electricity consumption using MLP

Figure 4.19 shows the forecasted commercial electricity consumption and its share of the total electricity consumption between the years 2021 and 2050. Where in 2021 reached 46072.68 GWh with a share of 17% and in 2050 went down to 42893.02 GWh with a share of 12.18%.

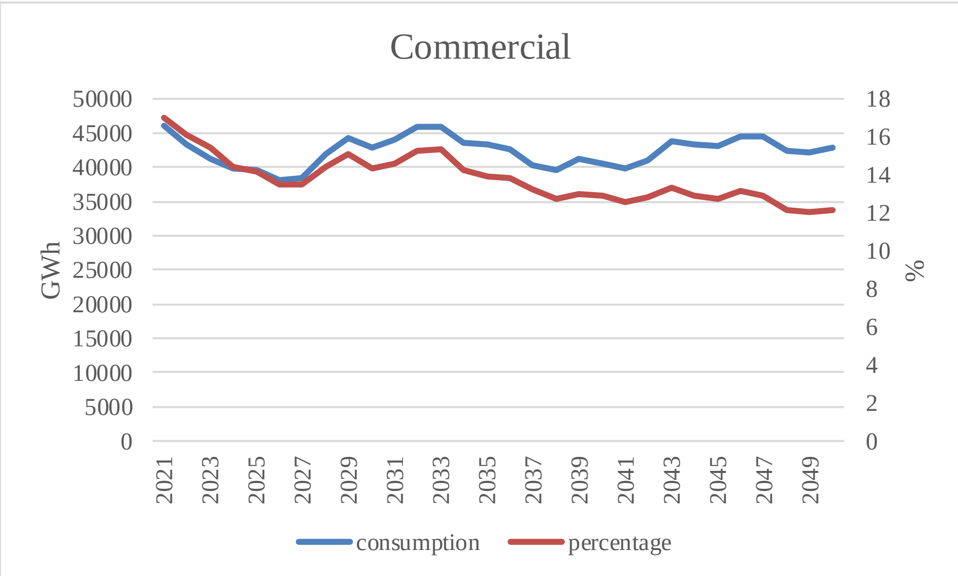

Figure 4.19 Forecasted commercial electricity consumption using MLP

The governmental sector went from 11724.03 GWh which made 4.32% of the total electricity consumption in 2021 to 16022.57 GWh which made 4.55% of the total electricity consumption in 2050 as presented in Figure 4.20.

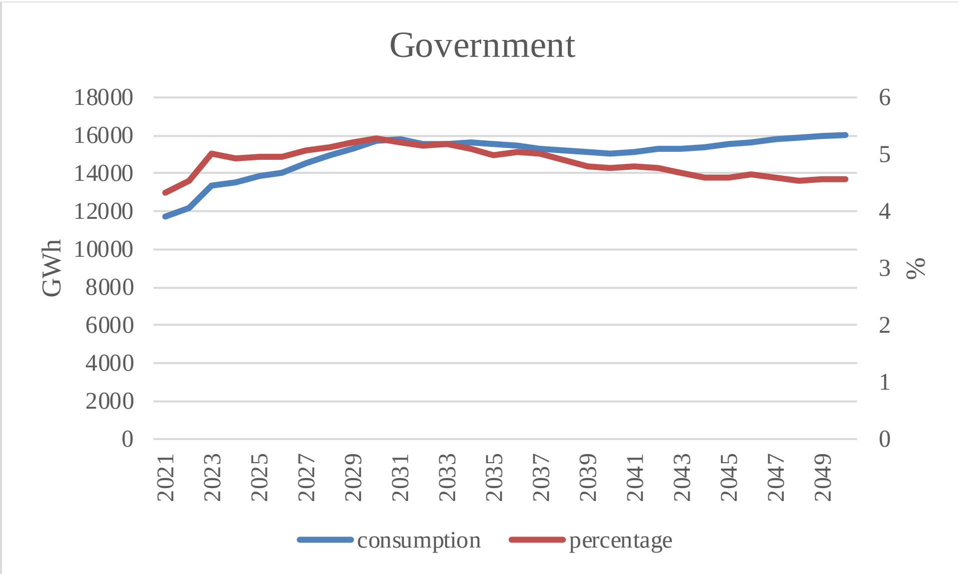

Figure 4.20 Forecasted government electricity consumption using MLP

The forecasted electricity consumption for the household sectors in 2021 was 61709.05 GWh that making 22.77% of the total electricity consumption whereas in 2050 it is forecasted to reach 91641.69 GWh with a share of 26.03% of the total electricity consumption as illustrated in Figure 4.21.

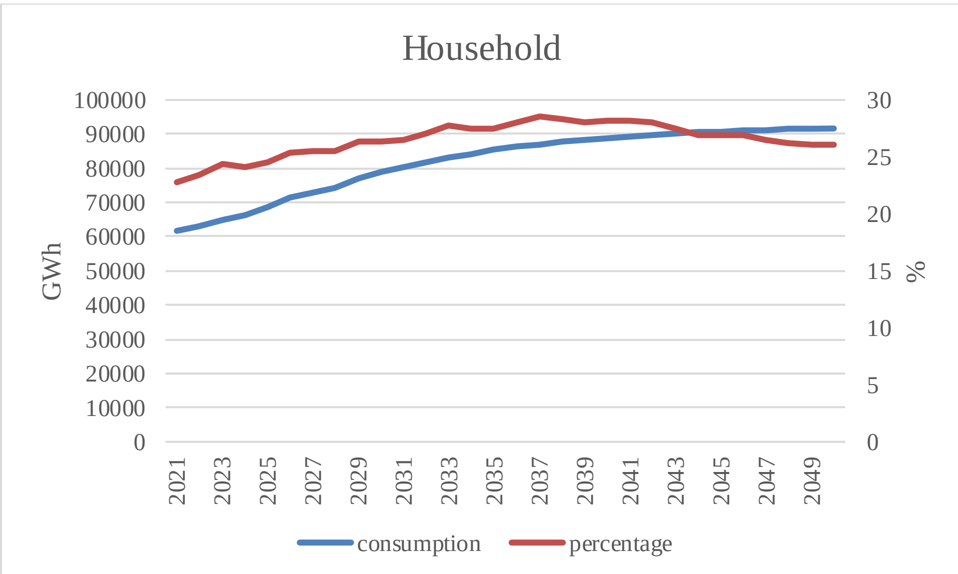

Figure 4.21 Forecasted household electricity consumption using MLP

Figure 4.22 shows the forecasted electricity consumption for the illumination sector which is 5247.7 GWh and 6805.97 GWh in 2021 and 2050 respectively. As the illumination sector kept almost the same share of the total electricity consumption as 1.93% through this period.

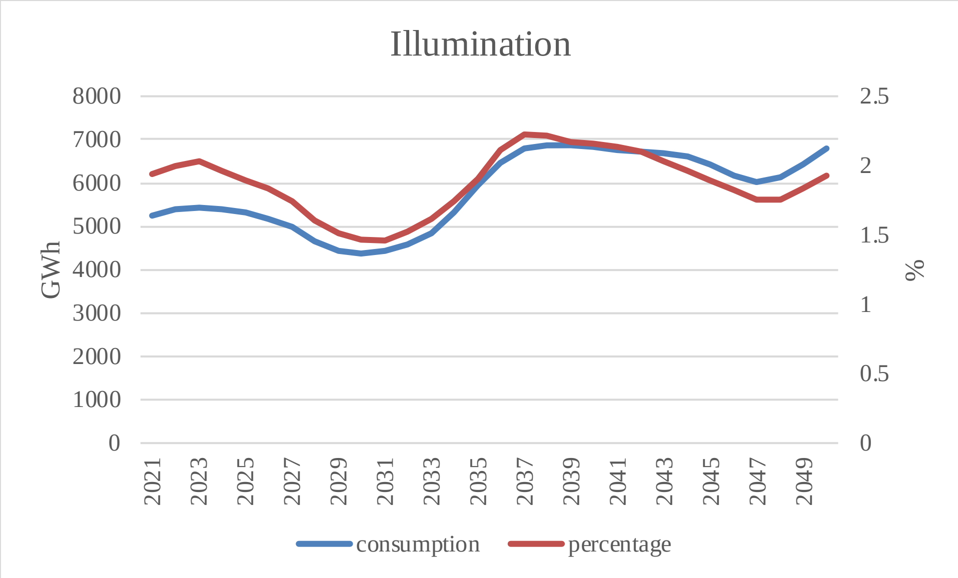

Figure 4.22 Forecasted illumination electricity consumption using MLP

The industrial sector has the biggest electricity consumption among other sectors where it is forecasted to consume 132069.44 GWh in 2021 and reach 170570.7 GWh in 2050. During this period the industrial sector's share of the total electricity consumption fluctuated where in 2021 it is forecasted to be 48.74% and stayed around this percentage during this time period as illustrated in the Figure 4.23.

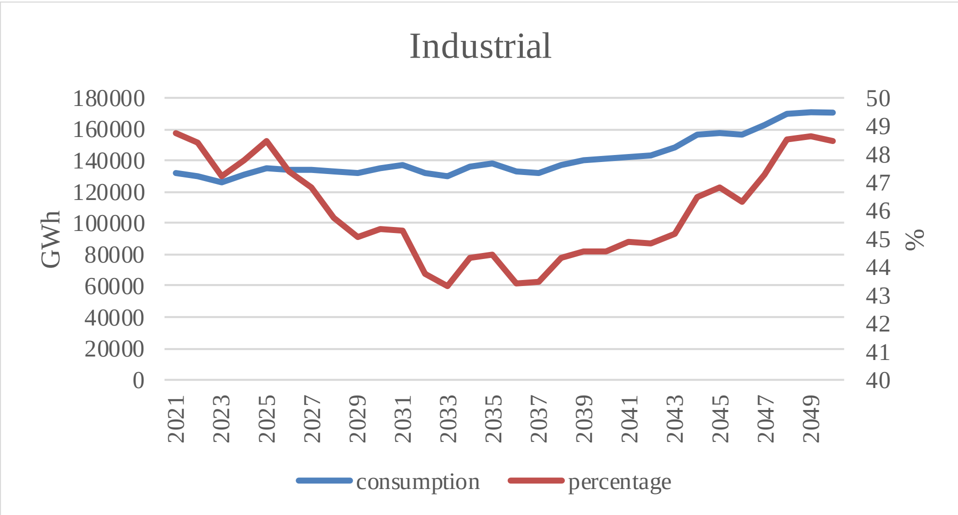

Figure 4.23 Forecasted industrial electricity consumption using MLP

The combined other sectors made about 5.2% of the total electricity consumption in 2021 and in 2050 reached 6.8 % as forecasted. The forecasted electricity consumption for the other sectors is deemed to be 14095.55 GWh and 24041.58 GWh in 2021 and 2050 respectively as shown in Figure 4.24.

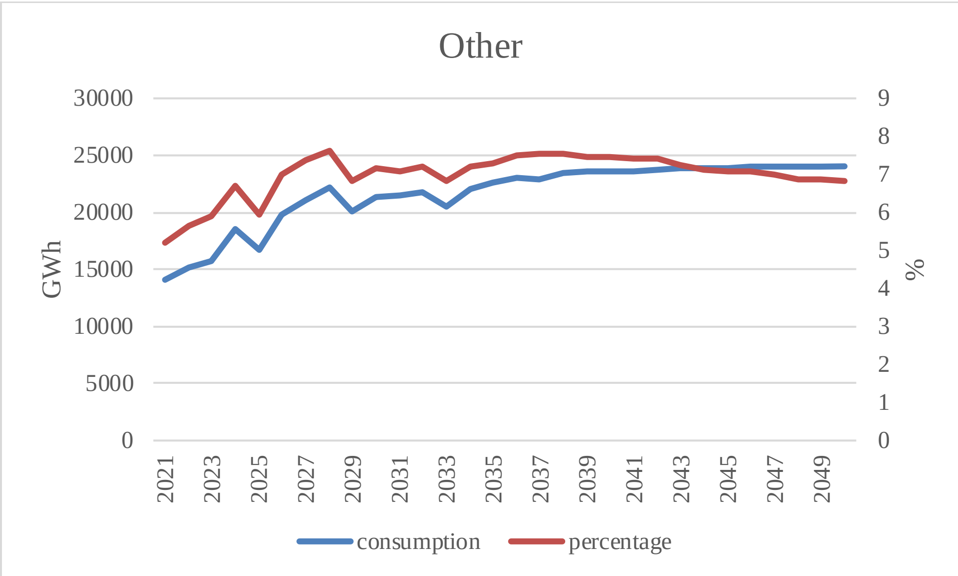

Figure 4.24 Forecasted other sectors' electricity consumption using MLP

The forecasted total electricity consumption is shown in Table 4.10.

Table 4.10 Forecasted total electricity consumption using MLPRegressor

Year	Total	Year	Total	Year	Total
2021	284056.7	2031	329585.4	2041	347206.8
2022	290339.7	2032	332039.9	2042	348399.2
2023	295984.2	2033	333549.3	2043	349561.8
2024	301661.6	2034	336054.1	2044	350661.8
2025	307099.6	2035	337941.3	2045	351725.6
2026	312243.1	2036	339729.7	2046	352683.2
2027	317493	2037	341438.8	2047	353616.3
2028	320734.3	2038	342999.9	2048	354489.3
2029	322824.8	2039	344428.1	2049	355314.4
2030	326311.3	2040	345844.5	2050	356092.2

4.6. Evaluation of ML algorithms

Different statistical metrics used to evaluate the models’ performance and the accuracy of the generated formula.

The forecasted data has 41 instances in order to compare the results of the evaluation metrics performed on the first ten instances this means the MAE, MAPE, RMSE, and MSE were calculated for the period between 2021-2030 but the determination coefficient R^2 was calculated for whole data.

Table 4.11 shows the evaluation metrics for the forecasted Commercial sector electricity consumption. For this attribute and considering the statistical metrics results, the MLPRegressor had better performance than SMOReg. where MAE, RMSE and MSE have a lower value and did not increase dramatically through the future instances and the MAPE kept going down through these instances for MLPRegressor method where it stayed under 10% and accordingly it is considered to have high accuracy, as for SMOReg the MAPE went up from less than 10% to reach 44.2% where the model deemed to have reasonable accuracy.

Table 4.11 Evaluation metrics for the forecasted commercial sector

Commercial								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	3412.07	7.62	5441.94	29614702.28	762.47	8.87	1333.52	1778281.52
2022	4718.65	10.40	7519.96	56549843.75	789.32	8.42	1354.82	1835546.34
2023	5384.18	11.60	8222.15	67603759.53	776.39	7.53	1341.47	1799540.54
2024	6240.26	13.37	8919.57	79558725.06	797.53	6.85	1361.52	1853726.00
2025	7295.31	15.61	9752.08	95103124.53	817.50	6.50	1411.15	1991332.06
2026	8299.59	17.63	10645.86	113334410.64	826.37	6.19	1418.82	2013063.93
2027	9317.37	19.61	11712.73	137187975.83	852.30	6.19	1459.02	2128748.08
2028	11594.54	24.47	13454.86	181033291.17	905.43	6.30	1575.90	2483463.54
2029	15648.33	33.39	16281.60	265090371.09	952.51	6.15	1621.23	2628383.13
2030	20145.27	44.30	20145.27	405831976.94	1013.28	6.13	1651.17	2726373.69

Table 4.12 shows the evaluation metrics for the forecasted Governmental sector electricity consumption. Both models had a good performance according to evaluation metrics. Where the MAPE for both is between 8-15% where this value gives the models good to high accuracy. but considering the rest of the metrics, MLPRegressor had lower and stable values overall.

Table 4.12 Evaluation metrics for the forecasted government sector

Government								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	856.50	8.52	980.29	960965.45	385.61	13.91	508.60	258673.74
2022	1073.51	10.51	1213.97	1473720.11	448.93	14.79	581.16	337749.71
2023	1062.47	10.11	1233.82	1522316.94	451.34	13.20	585.16	342416.15
2024	1263.09	11.49	1523.17	2320046.29	426.09	11.44	565.55	319842.85
2025	1364.09	11.84	1662.67	2764467.10	421.54	10.51	564.88	319086.00
2026	1272.94	10.28	1561.36	2437834.55	414.18	9.32	536.39	287712.64
2027	1369.85	10.66	1608.55	2587433.23	393.47	8.36	519.84	270236.86
2028	1772.79	13.69	1859.22	3456708.85	412.20	9.56	532.28	283323.38
2029	1951.32	14.60	2055.37	4224561.78	447.04	10.87	568.05	322685.40
2030	1341.90	10.47	1341.90	1800693.67	450.65	10.43	570.67	325665.06

Table 4.13 shows the evaluation metrics for the forecasted Household sector electricity consumption. SMOReg and MLPRegressor are considered to have high accuracy as both registered a value of less than 10% for MAPE. Even MLPRegressor method had better statistical results at first but SMOReg method results kept going down through the instances.

Table 4.13 Evaluation metrics for the forecasted household sector

Household								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	1979.91	4.62	2457.63	6039932.40	628.98	3.30	841.34	707858.07
2022	2375.11	5.44	2933.64	8606257.19	770.25	3.82	1039.81	1081212.29
2023	2079.79	4.67	2696.69	7272138.77	803.89	3.86	1071.72	1148574.42
2024	1742.44	3.78	2243.83	5034761.38	810.16	3.74	1076.20	1158202.46
2025	1911.38	3.99	2169.23	4705575.99	846.58	3.74	1111.64	1235743.88
2026	1830.20	3.74	2088.92	4363582.44	917.16	3.90	1195.39	1428966.90
2027	1358.62	2.71	1592.27	2535322.05	960.46	3.99	1247.62	1556554.72
2028	943.91	1.82	1075.84	1157441.14	965.49	3.83	1261.83	1592227.00
2029	971.40	1.85	1043.32	1088512.60	984.76	3.71	1265.13	1600565.47
2030	925.76	1.74	1059.66	1122888.63	1019.79	3.76	1296.88	1681889.33

Table 4.14 shows the evaluation metrics for the forecasted Illumination sector electricity consumption. MLPRegressor gave better results with higher accuracy than SMOReg for this attribute.

Table 4.14 Evaluation metrics for forecasted illumination sector

Illumination								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	396.11	8.52	505.79	255823.74	153.59	6.59	181.38	32898.49
2022	656.46	13.96	825.55	681533.49	196.11	8.09	226.46	51282.50
2023	725.42	15.36	876.70	768609.03	220.34	8.20	253.97	64499.66
2024	753.18	15.94	853.07	727720.53	234.62	7.84	268.74	72222.56
2025	801.89	16.81	875.62	766711.84	239.59	7.43	272.80	74417.58
2026	874.96	18.07	935.71	875548.22	236.04	7.16	267.72	71672.65
2027	1001.30	20.34	1035.26	1071763.59	234.46	7.01	264.40	69909.50
2028	1110.92	22.11	1131.44	1280151.81	253.83	7.48	277.88	77217.69
2029	1262.29	24.47	1263.70	1596935.89	274.74	7.81	294.14	86516.69
2030	1320.69	25.22	1320.69	1744225.08	283.48	7.90	303.54	92137.61

Table 4.15 shows the evaluation metrics for the forecasted Industrial sector electricity consumption. the accuracy for both methods was considered high as both have a MAPE of less than 10% but through the instances, SMOReg performed better where the value of the evaluation metrics kept declining where the lower the evaluation metrics the better the performance.

Table 4.15 Evaluation metrics for the forecasted industrial sector

Industrial								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	5730.80	5.52	6496.05	42198630.32	2995.48	4.62	3652.86	13343374.76
2022	4991.75	4.74	5839.37	34098287.54	3691.87	5.41	4547.21	20677163.72
2023	4801.37	4.45	5751.02	33074226.12	3299.85	4.54	4089.67	16725362.03
2024	4886.92	4.43	5988.55	35862716.82	2770.67	3.72	3494.38	12210704.23
2025	6195.92	5.49	7178.31	51528090.51	2831.44	3.76	3514.03	12348421.08
2026	5847.19	5.09	7084.45	50189421.17	2983.08	3.75	3767.57	14194599.68
2027	5510.00	4.70	7045.27	49635805.82	3240.12	3.98	4104.69	16848457.41
2028	3602.91	3.06	4806.12	23098827.27	3393.13	4.17	4207.84	17705933.27
2029	1635.95	1.39	1636.29	2677455.79	3512.10	4.34	4277.02	18292921.26
2030	1503.02	1.25	1503.02	2259079.86	3925.67	4.88	4638.17	21512658.01

Table 4.16 shows the evaluation metrics for the forecasted other sectors' electricity consumption. For this sector MLPRegressor performed better as the value for all the evaluation metrics is lower than the evaluation metrics registered by SMOReg.

Table 4.16 Evaluation metrics for the forecasted other sectors

Other								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	3122.41	23.64	3392.61	11509793.10	318.80	6.01	413.39	170887.84
2022	2526.49	19.52	2789.44	7780998.68	347.60	6.09	439.93	193536.72
2023	2664.54	20.56	2890.50	8354968.06	357.49	6.03	448.05	200747.94
2024	2227.49	16.44	2500.38	6251889.47	374.44	5.96	467.85	218886.56
2025	1676.49	11.86	2003.75	4015031.23	384.93	5.83	481.01	231370.77
2026	1542.15	9.98	1718.44	2953042.01	400.44	5.66	504.82	254842.99
2027	1509.85	9.66	1733.24	3004116.11	425.49	5.69	532.83	283908.73
2028	1070.04	6.19	1104.06	1218946.50	439.76	5.94	541.76	293498.89
2029	1097.29	6.26	1146.05	1313420.93	564.91	6.92	701.19	491672.19
2030	755.42	4.09	755.42	570655.34	597.28	7.15	746.87	557814.48

Table 4.17 shows the evaluation metrics for the forecasted total electricity consumption. Both ML algorithms had a high accuracy for forecasting total electricity consumption where the MAPE is less than 10% in addition to that the rest of the evaluation metrics have approximately close results.

Table 4.17 Evaluation metrics for the forecasted total electricity consumption

Total								
	SMOReg				MLPRegressor			
N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2021	5215.57	2.23	6435.39	41414280.64	3575.28	2.63	5321.07	28313792.13
2022	5049.14	2.10	6312.06	39842038.99	3892.95	2.78	5670.28	32152048.04
2023	5313.41	2.16	6714.69	45087092.28	3984.96	2.84	5735.74	32898666.38
2024	5890.75	2.37	7172.50	51444801.81	4017.63	2.83	5762.81	33209998.63
2025	6301.82	2.49	7626.79	58167992.99	4101.88	2.81	5839.07	34094739.90
2026	6596.22	2.56	7953.78	63262607.21	4207.52	2.87	5865.94	34409245.24
2027	8265.29	3.21	8925.41	79662954.56	4368.27	2.95	6049.82	36600307.22
2028	8707.02	3.35	9492.98	90116632.00	4492.56	3.02	6145.09	37762074.31
2029	10624.42	4.07	11088.51	122955098.43	4544.88	3.03	6182.65	38225182.12
2030	13866.40	5.28	13866.40	192277023.71	4887.74	3.19	6402.89	40997019.07

Table 4.18 presents the determination coefficient R^2 where it was calculated for each forecasted attribute as well as for the total forecasted electricity consumption which is the summation of the forecasted sectors presented in the table as Total S and the forecasted total electricity consumption presented in the table as Total F.

The R^2 results for MLPRegressor were higher than SMOReg for all the attributes. But both have a result higher than 75% this yields the forecasted data for both near the actual data.

Table 4.18 Determination coefficient

Attribute	SMOReg	MLPRegressro
Commercial	76.10%	84.05%
Government	88.47%	94.25%
Household	94.11%	97.09%
Illumination	80.87%	88.27%
Industrial	92.29%	96.79%
Other	93.62%	96.80%
Total S	87.77%	96.67%
Total F	91.43%	95.90%

Table 4.21 illustrates the APE for the performance of SMOReg model as the highest APE was recorded at 5.01% in 2020 and the lowest APE was recorded at 0.03% in 2016.

Table 4.19 The percentage error for SMOReg formula

Year	Actual	Forecasted	APE
2011	186100.0	181239.87	2.68%
2012	194923.0	190217.67	2.47%
2013	198045.0	199447.97	0.70%
2014	207375.0	210922.76	1.68%
2015	217312.0	220597.06	1.49%
2016	231203.7	231134.05	0.03%
2017	249022.6	242080.84	2.87%
2018	258232.0	253417.52	1.90%
2019	257273.1	264723.11	2.81%
2020	262702.13	276568.53	5.01%

Fig 4.25 depicts the actual and forecasted total electricity consumption that has been obtained by SMOReg formula.

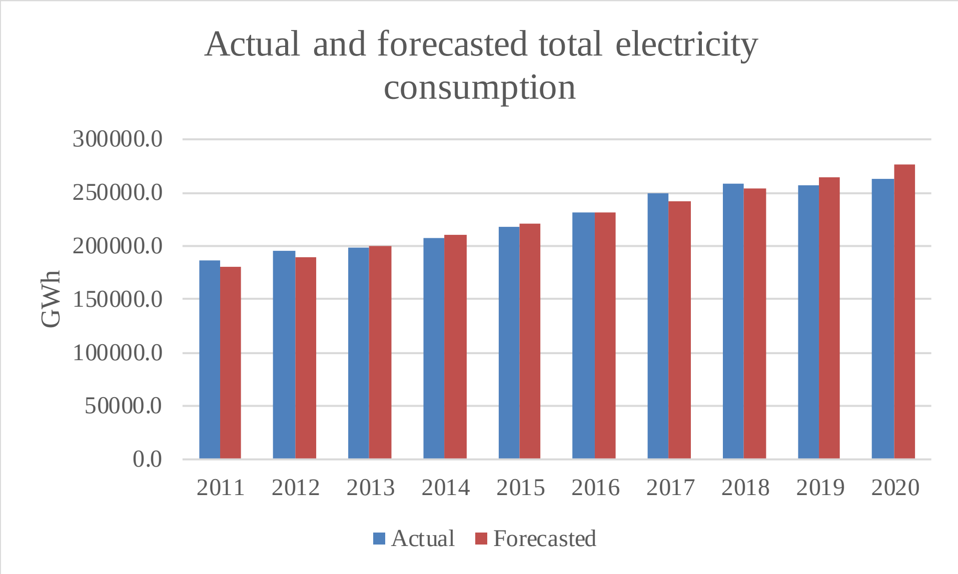

Figure 4.25 Actual and forecasted total electricity consumption with SMOReg

Fig 4.26 shows the total electricity consumption from 1970 until 2050 where the forecasted results used the MLPRegressor method.

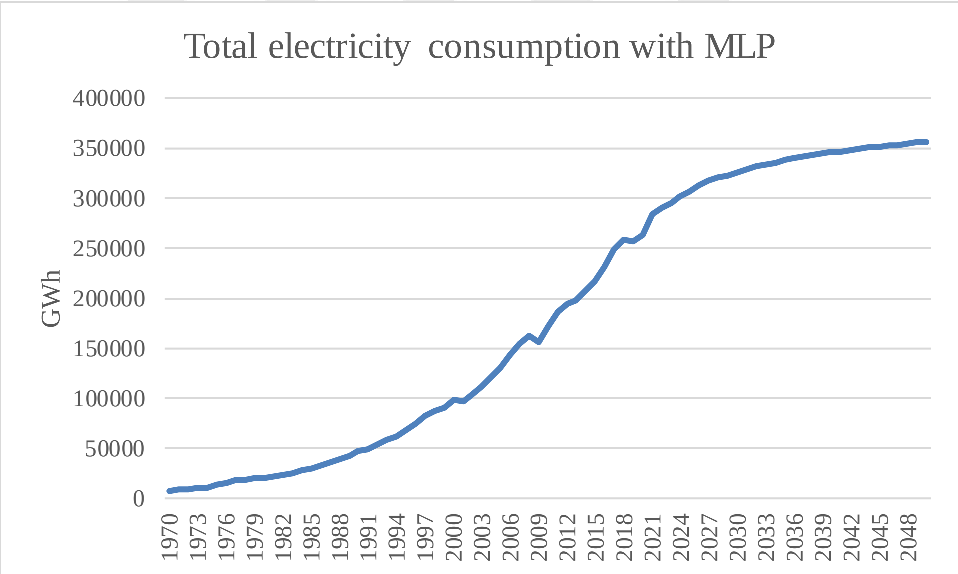

Figure 4.26 Total electricity consumption with MLP

Fig 4.27 shows the total electricity consumption from 1970 until 2050 as the forecasted results used SMOReg method.

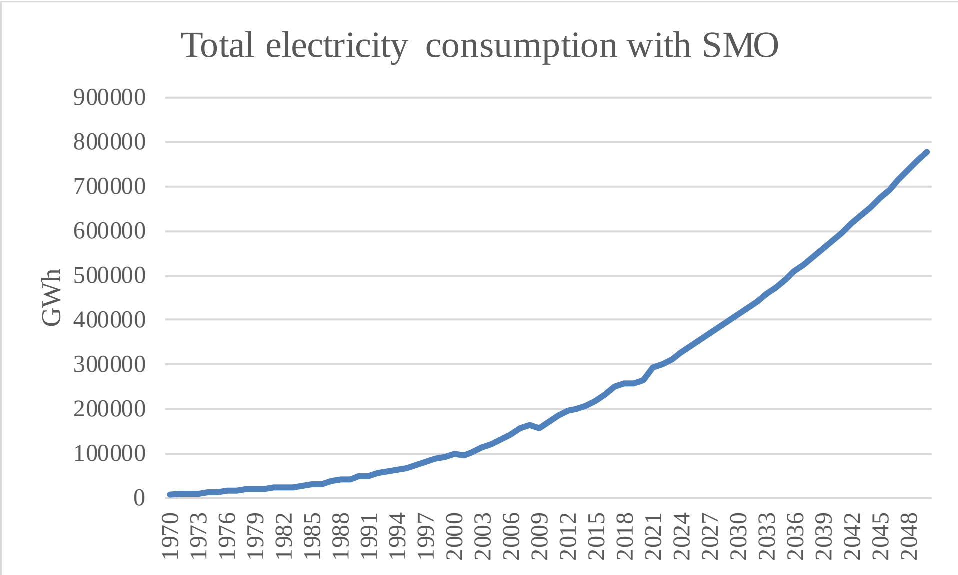

Figure 4.27 Total electricity consumption with SMO

Both methods performed very well according to the evaluation metrics and by comparing the output of both methods and analyzing the tendency of the forecasted value according to the actual tendency of the actual data from 1970 until 2020 it is found that SMOReg gave better results than MLPRgressor in forecasting Turkey's total electricity consumption for the period 2021-2050.

5. CONCLUSION

This paper presented the need to forecast the electricity consumption for Turkey and an overview of previous studies that used different forecasting methods. Sectoral and total electricity consumption in Turkey for the years 2021-2050 were forecasted using with MLPRegressor and SMOReg time series forecasting models. Different metrics were used to evaluate the performance of each model namely R^2, MAE, MAPE, MSE, and RMSE.

After evaluation, it is observed that for the commercial, governmental, illumination, and other sectors MLPRegressor performed better.

For forecasting the industrial and household sectors, SMOReg yeilded better performance with higher accuracy through the forecasted data.

For the total electricity consumption and after analyzing the trend that both models gave of the forecasted total electricity consumption, it is noticed that SMOReg gave better results according to the previous total electricity consumption trend.

By using SMOReg formula in forcasting the total electricity consumption for turkey in sectoral bases it is found that in the year 2021 the total electricity consumption wil be 292142 GWh and will reach 777854 GWh in the year 2050.

As a result of evaluating both models and in the scoop of forecasting Turkey’s total electricity consumption it is found that SMOReg is the better method for doing so. The results of this study can be used in helping Turkey's electricity production companies to make sure that the future electricity demand can be fulfilled so that the economy would keep getting better and grow faster.

BIBLIOGRAPHY

Abdel-Aal, R. E., & Al-Garni, A. Z. (1997). *FORECASTING MONTHLY ELECTRIC ENERGY CONSUMPTION IN EASTERN SAUDI ARABIA USING UNIVARIATE TIME-SERIES ANALYSIS* (Vol. 22).

Ağbulut, Ü. (2022). Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms. *Sustainable Production and Consumption*, *29*, 141–157. https://doi.org/10.1016/j.spc.2021.10.001

Ağbulut, Ü., Gürel, A. E., & Biçen, Y. (2021). Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, *135*. https://doi.org/10.1016/j.rser.2020.110114

Ahmad, A. S., Hassan, M. Y., Abdullah, M. P., Rahman, H. A., Hussin, F., Abdullah, H., & Saidur, R. (2014). A review on applications of ANN and SVM for building electrical energy consumption forecasting. In *Renewable and Sustainable Energy Reviews* (Vol. 33, pp. 102–109). Elsevier Ltd. https://doi.org/10.1016/j.rser.2014.01.069

al Miraz, K. F., Biswas, K., & Rudra, R. Y. (2021). *Forecasting Electricity Demand of Bangladesh and its relation with physical and natural variables using group method of data handling model*.

Becirovic, E., & Cosovic, M. (2016, November 18). Machine learning techniques for short-Term load forecasting. *4th International Symposium on Environment Friendly Energies and Applications, EFEA 2016*. https://doi.org/10.1109/EFEA.2016.7748789

Bilal, M., Fayaz, M., Kim, H., & Pawar, P. (2020). *Comparative Analysis of Time Series Forecasting Approaches for Household Electricity Consumption Prediction*. https://www.researchgate.net/publication/355481783

Bilgili, M. (2009). ESTIMATION OF NET ELECTRICITY CONSUMPTION OF TURKEY. *J. of Thermal Science and Technology*, *29*, 89–98.

Ceylan, H., & Ozturk, H. K. (2004). Estimating energy demand of Turkey based on economic indicators using genetic algorithm approach. *Energy Conversion and Management*, *45*(15–16), 2525–2537. https://doi.org/10.1016/j.enconman.2003.11.010

Ceylan, Z. (2020). Assessment of agricultural energy consumption of Turkey by MLR and Bayesian optimized SVR and GPR models. *Journal of Forecasting*, *39*(6), 944–956. https://doi.org/10.1002/for.2673

Ceylan, Z. (2021). The impact of COVID-19 on the electricity demand: a case study for Turkey. *International Journal of Energy Research*, *45*(9), 13022–13039. https://doi.org/10.1002/er.6631

Ceylan, Z., Akbulut, D., & Baytürk, E. (2022). Forecasting gasoline consumption using machine learning algorithms during COVID-19 pandemic. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*. https://doi.org/10.1080/15567036.2021.2024919

Ceylan, Z., & Atalan, A. (2021). Estimation of healthcare expenditure per capita of Turkey using artificial intelligence techniques with genetic algorithm-based feature selection. *Journal of Forecasting*, *40*(2), 279–290. https://doi.org/10.1002/for.2747

Chou, J. S., & Tran, D. S. (2018). Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. *Energy*, *165*, 709–726. https://doi.org/10.1016/j.energy.2018.09.144

COMERT, M., & YILDIZ, A. (2021). A novel artificial neural network model for forecasting electricity demand enhanced with population-weighted temperature mean and the unemployment rate. *Turkish Journal of Engineering*. https://doi.org/10.31127/tuje.903876

Cortes, C., Vapnik, V., & Saitta, L. (1995). Support-Vector Networks Editor. In *Machine Leaming* (Vol. 20). Kluwer Academic Publishers.

Csépe, Z., Makra, L., Voukantsis, D., Matyasovszky, I., Tusnády, G., Karatzas, K., & Thibaudon, M. (2014). Predicting daily ragweed pollen concentrations using Computational Intelligence techniques over two heavily polluted areas in

Europe. *Science of the Total Environment*, *476–477*, 542–552. https://doi.org/10.1016/j.scitotenv.2014.01.056

David, T. W., Anizelli, H., Jacobsson, T. J., Gray, C., Teahan, W., & Kettle, J. (2020). Enhancing the stability of organic photovoltaics through machine learning. *Nano Energy*, *78*. https://doi.org/10.1016/j.nanoen.2020.105342

Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. In *Renewable and Sustainable Energy Reviews* (Vol. 74, pp. 902–924). Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.02.085

Ediger, V. S. ß, & Useyin Tatlõdil, H. (2002). *Forecasting the primary energy demand in Turkey and analysis of cyclic patterns*. www.elsevier.com/locate/enconman

Feng, X., Ma, G., Su, S. F., Huang, C., Boswell, M. K., & Xue, P. (2020). A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan. *Ocean Engineering*, *211*. https://doi.org/10.1016/j.oceaneng.2020.107526

Geliştirme, S., & Başkanliği, D. (2021). *TÜRKİYE ELEKTRİK DAĞITIM A.Ş. GENEL MÜDÜRLÜĞÜ 2020 YILI TÜRKİYE ELEKTRİK DAĞITIMI SEKTÖR RAPORU*. https://www.tedas.gov.tr/sx.web.docs/tedas/docs/Stratejikplan/2020_Yili_Turkiye_Elektrik_Dagitimi_Sektor_Raporu.pdf

Hamzaçebi, C. (2007). Forecasting of Turkey's net electricity energy consumption on sectoral bases. *Energy Policy*, *35*(3), 2009–2016. https://doi.org/10.1016/j.enpol.2006.03.014

Hong, W.-C., Dong, Y., Lai, C.-Y., Chen, L.-Y., & Wei, S.-Y. (2011). *SVR with Hybrid Chaotic Immune Algorithm for Seasonal Load Demand Forecasting*. *4*, 960–977. https://doi.org/10.3390/en4060960

Jain, R., Nayyar, A., Arora, S., & Gupta, A. (2021). A comprehensive analysis and prediction of earthquake magnitude based on position and depth parameters using machine and deep learning models. *Multimedia Tools and Applications*, *80*(18), 28419–28438. https://doi.org/10.1007/s11042-021-11001-z

Kandananond, K. (2011). Forecasting electricity demand in Thailand with an artificial neural network approach. *Energies*, *4*(8), 1246–1257. https://doi.org/10.3390/en4081246

Kavitha, S., Varuna, S., & Ramya, R. (2017, May 1). A comparative analysis on linear regression and support vector regression. *Proceedings of 2016 Online International Conference on Green Engineering and Technologies, IC-GET 2016*. https://doi.org/10.1109/GET.2016.7916627

Kaytez, F., Taplamacioglu, M. C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power and Energy Systems*, *67*, 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036

Kermanshahi, B., & Iwamiya, H. (2002). *Up to year 2020 load forecasting using neural nets*. www.elsevier.com/locate/ijepes

Lydia, M., & Edwin Prem Kumar, G. (2022). Soft computing models for forecasting day-ahead energy consumption. *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2022.03.003

Nwulu, N. I., & Agboola, O. P. (2012). Modelling and predicting electricity consumption using artificial neural networks. *2012 11th International Conference on Environment and Electrical Engineering, EEEIC 2012 - Conference Proceedings*, 1059–1063. https://doi.org/10.1109/EEEIC.2012.6221536

OECD (2022), Gross domestic product (GDP) (indicator). doi: 10.1787/dc2f7aec-en (Accessed on 10 June 2022)

Oğcu, G., Demirel, O. F., & Zaim, S. (2012). Forecasting Electricity Consumption with Neural Networks and Support Vector Regression. *Procedia - Social and Behavioral Sciences*, *58*, 1576–1585. https://doi.org/10.1016/j.sbspro.2012.09.1144

Saab, S., Badr, E., & Nasr, G. (2001). Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. In *Energy* (Vol. 26). www.elsevier.com/locate/energy

Shevade, S. K., Keerthi, S. S., Bhattacharyya, C., & Murthy, K. R. K. (2000). Improvements to the SMO Algorithm for SVM Regression. In *IEEE TRANSACTIONS ON NEURAL NETWORKS* (Vol. 11, Issue 5). https://doi.org/10.1109/72.870050

Smola, A. J., Sch¨olkopf, B., & Sch¨olkopf, S. (2004). A tutorial on support vector regression*. In *Statistics and Computing* (Vol. 14). Kluwer Academic Publishers.

Taghi Sattari, M., Feizi, H., Samadianfard, S., Falsafian, K., & Salwana, E. (2021). Estimation of monthly and seasonal precipitation: A comparative study using data-driven methods versus hybrid approach. *Measurement: Journal of the International Measurement Confederation*, *173*. https://doi.org/10.1016/j.measurement.2020.108512

Tunç, M., Çamdali, Ü., & Parmaksizoglu, C. (2006). Comparison of Turkey's electrical energy consumption and production with some European countries and optimization of future electrical power supply investments in Turkey. *Energy Policy*, *34*(1), 50–59. https://doi.org/10.1016/j.enpol.2004.04.027

TurkStat - Data Portal. (2022). Retrieved May 20, 2022, from https://data.tuik.gov.tr/Kategori/GetKategori?p=cevre-ve-enerji-103&dil=2

Usha, T. M., & Balamurugan, S. A. A. (2016). Seasonal Based Electricity Demand Forecasting Using Time Series Analysis. *Circuits and Systems*, *07*(10), 3320–3328. https://doi.org/10.4236/cs.2016.710283

Velasquez, C. E., Zocatelli, M., Estanislau, F. B. G. L., & Castro, V. F. (2022). Analysis of time series models for Brazilian electricity demand forecasting. *Energy*, *247*. https://doi.org/10.1016/j.energy.2022.123483

Weka 3 - Machine Learning Software in Java. (n.d.). Retrieved May 20, 2022, from https://www.cs.waikato.ac.nz/ml/weka/

Yalcinoz, T., & Eminoglu, U. (2005). Short term and medium term power distribution load forecasting by neural networks. *Energy Conversion and Management*, *46*(9–10), 1393–1405. https://doi.org/10.1016/j.enconman.2004.07.005

APPENDIX: GENERATED FORMULAS FOR SMOReg MODEL

For the commercial sector, the forecasting formula is:

+ 0.2053 * (normalized) year-remapped

+ 0.3144 * (normalized) Lag_commercial-1

+ 0.0159 * (normalized) Lag_commercial-2

- 0.0198 * (normalized) Lag_commercial-3

+ 0.0879 * (normalized) Lag_commercial-4

- 0.712 * (normalized) year-remapped^2

+ 1.0277 * (normalized) year-remapped^3

+ 0.179 * (normalized) year-remapped*Lag_commercial-1

- 0.0126 * (normalized) year-remapped*Lag_commercial-2

+ 0.0325 * (normalized) year-remapped*Lag_commercial-3

- 0.1125 * (normalized) year-remapped*Lag_commercial-4

- 0.0066

For the government sector, the forecasting formula is:

- 0.3161 * (normalized) year-remapped

+ 0.3388 * (normalized) Lag_government-1

+ 0.2046 * (normalized) Lag_government-2

+ 0.4934 * (normalized) Lag_government-3

- 0.1362 * (normalized) Lag_government-4

+ 0.5774 * (normalized) year-remapped^2

+ 0.9986 * (normalized) year-remapped^3

- 0.0257 * (normalized) year-remapped*Lag_government-1

- 0.6343 * (normalized) year-remapped*Lag_government-2

- 0.6237 * (normalized) year-remapped*Lag_government-3

+ 0.045 * (normalized) year-remapped*Lag_government-4

+ 0.0563

For the household sector, the forecasting formula is:

- 0.1039 * (normalized) year-remapped

+ 0.31 * (normalized) Lag_household-1

+ 0.3038 * (normalized) Lag_household-2

+ 0.3102 * (normalized) Lag_household-3

+ 0.0685 * (normalized) Lag_household-4

- 0.073 * (normalized) year-remapped^2

+ 1.2262 * (normalized) year-remapped^3

- 0.1125 * (normalized) year-remapped*Lag_household-1

- 0.4067 * (normalized) year-remapped*Lag_household-2

- 0.381 * (normalized) year-remapped*Lag_household-3

- 0.2185 * (normalized) year-remapped*Lag_household-4

+ 0.0527

For the illumination sector, the forecasting formula is:

- 0.1432 * (normalized) year-remapped

+ 1.0969 * (normalized) Lag_illumination-1

+ 0.8713 * (normalized) Lag_illumination-2

+ 0.3997 * (normalized) year-remapped^2

+ 0.1172 * (normalized) year-remapped^3

- 0.2677 * (normalized) year-remapped*Lag_illumination-1

- 0.9961 * (normalized) year-remapped*Lag_illumination-2

+ 0.01

For the industrial sector, the forecasting formula is:

+ 0.4082 * (normalized) year-remapped

\+ 0.3233 * (normalized) Lag_industrial-1

\+ 0.0591 * (normalized) Lag_industrial-2

\- 0.0261 * (normalized) Lag_industrial-3

\- 0.4097 * (normalized) Lag_industrial-4

\+ 0.3791 * (normalized) year-remapped^2

\+ 0.1732 * (normalized) year-remapped^3

\- 0.4037 * (normalized) year-remapped*Lag_industrial-1

\- 0.0682 * (normalized) year-remapped*Lag_industrial-2

\+ 0.0426 * (normalized) year-remapped*Lag_industrial-3

\+ 0.5386 * (normalized) year-remapped*Lag_industrial-4

\- 0.0121

For the other sectors, the forecasting formula is:

\- 0.3056 * (normalized) year-remapped

\+ 0.5369 * (normalized) Lag_other-1

\+ 0.0574 * (normalized) Lag_other-2

\+ 0.1233 * (normalized) Lag_other-3

\+ 0.1175 * (normalized) Lag_other-4

\+ 0.7201 * (normalized) year-remapped^2

\+ 0.8862 * (normalized) year-remapped^3

\- 0.6886 * (normalized) year-remapped*Lag_other-1

\- 0.1157 * (normalized) year-remapped*Lag_other-2

\- 0.1759 * (normalized) year-remapped*Lag_other-3

\- 0.198 * (normalized) year-remapped*Lag_other-4

\+ 0.047

For the total electricity consumption, the forecasting formula is:

\- 0.045 * (normalized) year-remapped

+ 0.2384 * (normalized) Lag_total-1

+ 0.1178 * (normalized) Lag_total-2

+ 0.0183 * (normalized) Lag_total-3

- 0.1523 * (normalized) Lag_total-4

- 0.0886 * (normalized) Lag_total-5

+ 0.6274 * (normalized) year-remapped^2

+ 0.556 * (normalized) year-remapped^3

- 0.3121 * (normalized) year-remapped*Lag_total-1

- 0.17 * (normalized) year-remapped*Lag_total-2

- 0.034 * (normalized) year-remapped*Lag_total-3

+ 0.1844 * (normalized) year-remapped*Lag_total-4

+ 0.027 * (normalized) year-remapped*Lag_total-5

+ 0.0389