TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Shan Ali Malik 195662IASM

Application of Artificial Neural Networks for Short Term Load Forecasting on Example of Publicly Available Estonian Data

Master's thesis

Supervisor:	Eduard Petlenkov
	PhD
Co-Supervisor	Vjatšeslav Škiparev

Tallinn 2021

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvarateaduse instituut

Shan Ali Malik 195662IASM

Tehisnärvivõrkude rakendamine elektrienergia tarbimise lühiajaliseks ennustamiseks Eesti kogutarbimise näitel

Magistritöö

Juhendaja: Eduard Petlenkov
PhD
Kaasjuhendaja Vjatšeslav Škiparev

Tallinn 2021

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Shan Ali Malik

10.05.2021

Abstract

In this dissertation, author estimated the efficiency of the forecast short-term electrical consumption of NordPool AS clients in Estonia by utilizing Artificial Neural Network and implement a user-friendly program for ANN application. Accurate forecasting allows utility companies to manage their power plant resources and demand of consumption of electricity. In addition to this, a literature review is done, including all the short-term load forecasting techniques used by different research groups starting from simple time-series approaches to the final Artificial Intelligence techniques.

This dissertation is mainly concerned with load forecasting with high forecasting error; many experiments are made to overcome this problem and suggested solutions and finding the most valuable structure of a model and suitable settings. As a result, very significant improvement in forecasting accuracy achieved. Secondly, K-fold cross-validation is performed to analyze the forecasting capabilities of the model deeply.

As a result, the forecast error is lowered to 3.77% as an average. In order to avoid retraining of the Artificial Neural Network model, a user-friendly program is made in MATLAB.

This research work is written in English and 81 pages long, including 21 figures and 23 tables.

List of abbreviations and terms

STLF	Short Term Load Forecasting
SVM	Support Vector Machine
AI	Artificial Intelligence
ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average
SARIMA	Seasonal Autoregressive Integrated Moving Average
GUI	Graphical User Interface
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
ReLU	Rectified Linear Unit
SVR	Support Vector Regression
UI	User Interface
WRLS	Weight Recursive Least Square
VFF	Variable Forgetting Factor
FNN	Fuzzy Neural Network
FL	Fuzzy Logic

Table of contents

List of figures

List of tables

1 Introduction

In today's world, energy is playing a vital role in our lives. There are several different forms of energy that we are using nowadays [1] i.e., potential energy obtained with respect to position of a body, kinetic energy obtained from motion, thermal energy obtained from heat, electrical energy from hydropower plants and many other generation sources, chemical energy that we obtain from fossil fuels, Nuclear energy obtained from the nucleus of atom when nucleus combine with each other or break off. wind energy from wind power turbines, and solar energy from sun, etc. Most of the energy is being utilized for residential, commercial, and industrial, and agriculture purposes. There are fundamentally two types of energy sources; one is a renewable energy source, and the other is a non-renewable energy source, these both sources has great impact on economic growth of a country [2]. Electricity is one form of energy generated by the motion of electrons from the cathode towards the anode in any conducting material. Electricity encompasses different areas like technology as well as research, production, consumption, and transmission. Globally it seems that the major source of electricity generation is from coal-fired power plants. It is followed by electricity generation by gas, hydropower plants, and nuclear power plants, making a lot of contribution in supplying electricity demand [3].

Electricity generating market in Europe is moving more towards renewable energy sources, electricity through wind generation rise to 9% and contribution of solar power plants electricity generation rose to 15% in 2020 [4]. Estonia, which is a small country in the Baltic region of northern Europe, a large amount of electricity is obtained from coal, natural gas, biofuel and waste, furnace oil, and very small amount of electricity is obtained from hydropower plants, solar and wind; In all of the energy sources coal is contributing a huge amount that is 71% of the total. Which is followed by biofuel and waste, that contributes 19% of electricity generation, and all the remaining sources of electricity generation are less than 10%[5].

Electricity streams through the veins of autonomous machines in industries that are the backbone of the import and export of the country and building the country's economy. Electricity is providing heating and cooling for residential and commercial spaces depends upon the geographical locations. Electricity is as important in our life as like water and food. As automation is increasing day by day, we move towards a more advanced and globally connected world. Adequate and satisfactory electricity generation is very important for sustainable economic development and maintaining all the world's luxuries.

To generate an adequate amount of electricity and provide the users continuous supply, we need to forecast electricity consumption with the least possible error. As researchers started researching this topic a long ago, and there are so many developments in this area. A brief survey has been done by Mohammad Nazeeruddin and Hesham K. Alfares [6] on all the relevant techniques and proposed methods and classification for energy forecasting.

Mainly there are three types of load forecasting techniques.

- Short term Load Forecast
- Medium-term Load forecast
- Long term Load forecasting

The most familiar type among all forecasting techniques is short-term forecasting, which is the electricity or load demand forecast from one hour to one day and from one day to one week. Power supply companies then use this forecasting data to analyse the amount of electricity needed for distribution to optimize the power production at power plants according to the load demand. Meanwhile, they will set up resources like purchasing oil, transporting coal, diesel as per requirements.

Load forecasting is not an easy task because of the dynamic behaviour of load consumption. It cannot be predicted exactly because of the uncertainty of load requirement on a specific day, but it can be reduced by decreasing the forecasting error [7]. Forecasting error is the difference between the actual values and the predicted values of the load. For Forecasting electrical consumption, there are several techniques and suggested methods. If we take ten residential buildings under consideration for forecasting the load and we keep a check on the activities of all peoples when they are at home and when they are not at home, and make our prediction for consumption of electricity, but somehow if some peoples will not go to work, as it happened in uncertain corona situation, peoples started working from home, so our forecast error will be more. Electricity forecasting depends upon several factors like season, weather conditions [8], geographical location, and type of energy zone like industrial, residential, commercial, or agriculture, etc.

1.1 Background

Estonia is a northern country, with a 1.325 million population and a total area of 42390km. sq. The majority of the population lives in cities, around 70% of the total population [9]. Estonian weather is very severe in winters; sometimes it goes around -25 degrees Celsius and followed warm summers, with average temperature, 18 degrees Celsius. Electricity consumption in Estonia shows a different pattern in winters and summers. There are many electricity suppliers in Estonia; starting from 2013 onward, all consumers have their own choice to select a particular supplier[10] All these suppliers have specific demand for electricity from their different consumer's class, and they generate and transmit electricity as per requirement. State own company Elering manages all these power plants. Elering is the foundation of the energy system in Estonia, which is ensuring the supply of electricity in the whole country without any failure. Nord Pool is the dominant power distribution market in Europe, which has many customers in Estonia[11]. Nord Pool is the place where the electricity is traded between consumers and producers, and they also provide forecasting of electricity prices.

1.2 Problem Statement

A common problem to deal with the electricity is to consume it immediately because we cannot afford energy storage for a whole country, therefore it is necessary to forecast the required energy ahead before production. The forecasting of electricity should be accurate enough, so that supplier will manage all the resources and order this demand of electricity a day before, to the companies generating electricity [11].

If the electrical load increases than the usual demand of electrical consumption, suppliers must increase their generating capacity and turn on more power plant stations that are usually standby power plants to meet the accidental load situation (peak load demand) or any uncertain increase in power consumption. Standby power plants operate when the electricity consumption suddenly increases during peak load hours. The standby power plants are diesel generators or furnace oil that turns on within seconds and starts generating power and meeting the load demand as per requirement. On the other side, there are baseload power plants, e.g., coal power plant that takes around 6 to 7 hour in a start up, so they can only provide a constant amount of fixed electricity for a week a month. Standby power plants compensate for any fluctuations in peak hours. Accordingly, when these standby power plants operate, the prices of electricity don't remain the same. It increases the normal generating cost of electricity almost two times because furnace-oil/Diesel prices are quite high and always change with the fluctuations in international oil market prices.

As electrical load is dependent on several factors in which some are nonlinear, like weather conditions due to global warming, we cannot predict exactly how climatic conditions will be changed in upcoming months. If the weather is hot and the humidity level increases, the electricity consumption will increase due to the excessive usage of air conditioners to maintain a good temperature. Likewise, if the weather is cold and less humid, then the heating demand will be increased. Therefore, in each case, we are dependent on electricity, and accurate forecasting is really important.

One example is considered here to explain how prices of electricity increases when standby power plants operates rather than base load power plants. Last year there was an increase in the per-unit cost of electricity due to a failure of a submarine electric cable, NordBalt, which connects Lithuania, Sweden, and Estonia. NordBalt is the power cable with an electrical capacity of 700 megawatts, which is the total consumption of Estonia in June. NordBalt is the only source of electric power generated by the hydropower plants, and it supplies to the Nordic countries[12]. This caused an astronomical rise in electricity prices to 255€ per megawatt-hour that day in Estonia. The government official’s bodies announced that the electricity prices during the peak hours in Estonia would be eight times higher than usual.

This research work is mainly concerned about the short-term load forecasting of electrical energy for NordPool-Estonia customers. It comprises all the variables that play a vital role in deciding the power consumption in the area. If some variable has a direct relation with the forecast and somehow is not given proper consideration in forecasting, then the forecasting error will be high.

Another problem is that the forecasted temperature is also not 100% exact compared to the actual temperature of the day. If the forecasted temperature value varies, that will also decrease the forecasting accuracy. Secondly, due to the inclusion of the global warming effect and the nonlinear behaviour of energy consumption, the data collected from the previous year is not that valuable for forecasting energy in the future. Last year in 2020 was the warmest year in the history of Estonia[13]; therefore, electrical consumption in march 2020 was different from march 2021.

In this research work, the data has been collected through various resources and the companies predicting weather conditions. After that, statistical analysis was performed and processed to train the artificial neural network through AI techniques. Finally, the data is being analysed to forecast the energy consumption for short-term load forecasting in Estonia.

1.3 Aim and Objective

This research is concerned with electrical load forecasting through artificial intelligence techniques of STLF. Many artificial intelligence methods are commonly proposed for short-term load forecasting. This dissertation mainly focused on artificial neural network methods, Support Vector Machine SVM [14]; meanwhile, the drawback of SVM is that it can only be used for a small number of data sets, and its results are not very transparent [14]. On the other hand, the neural network is very prominent in this regard, and their accuracy increases as the amount of learning data set increases. Therefore, more preference in this research work is granted to the artificial neural network.

1.3.1 Objectives

- Look into MATLAB neural network tool for short-term load forecasting (STLF) and examine their accuracy.

- Overview different approaches for STLF and make a conclusive model for state-of-the-Art techniques.

- To collect the energy consumption data set of inputs from an electrical power supplier NordPool and the temperature data from the website rp5.ru [15]. Which regulates and forecasts the weather.

- Performed experiments on neural networks by training them with a small amount of data from one day, three months, then finally one year, using different layers, different number of epochs, and input parameters to get the best and more precise forecasted results.

- Reduce the forecasting error in STLF.

- Developing an open-source, user-friendly MATLAB program to estimate the forecast in the future.

- To validate the results with the actual power consumption data provided by Nordpool[16].

- Summarizing the results in the form of a technical report and submitting it to the department of computer system engineering TalTech.

1.4 Work Flow of the thesis

This thesis is based on a real-world problem in the electricity market that is load forecasting which is not just important for electricity consumers. Still, it is also very important for utility companies.

There is a lot of research on this real-world problem for load forecasting, using the time series techniques and AI methods. However, this thesis is prioritized by investigating electrical forecasting using neural networks that are still on top with the best accuracy. Therefore, the first chapter of this research starts with an introduction followed by a deep study of the literature review.

In the third stage of this thesis, the Methodology is discussed, MATLAB tools, and finally experiments. The 4th stage is based on analysis of the obtained experimental results and finally conclude them.

At final stage in chapter 5, there is a comprehensive discussion of all experimental results and in last there is a conclusion of this research work. General work flow of this thesis is shown in figure 1.

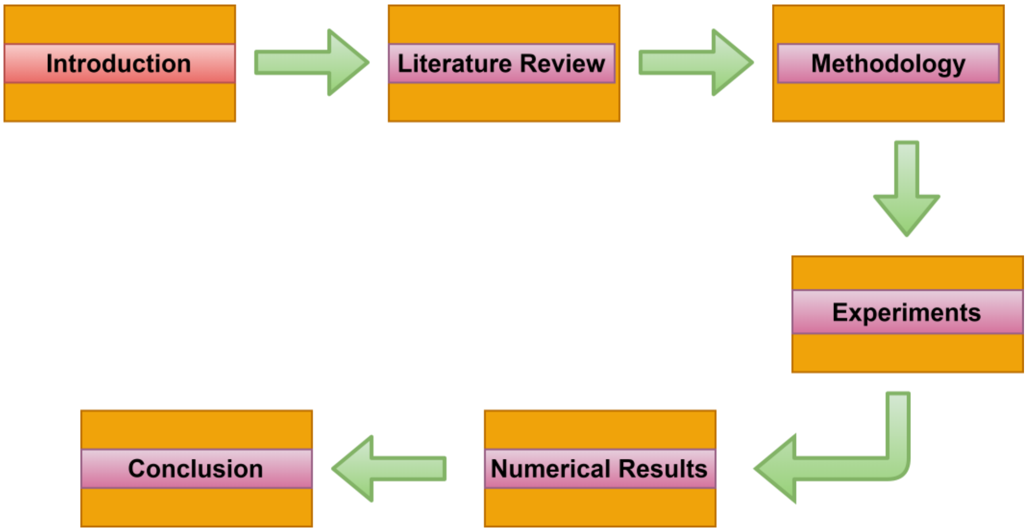

Figure 1. Work flow of ANN STLF research work.

2 State of the Art methods

Load forecasting started in 1987, and by reviewing different researches in history, the first forecasting was based on time series methods which include Autoregression (AR), Moving Average (MA), ARMA, ARIMA, and SARIMA, etc. Later it was observed that machine learning techniques show much better performance, so all research diverted towards artificial intelligence(AI). Artificial intelligence has shown very good performance, and still, this topic is a hot favourite in forecasting electrical consumption.

At this stage, different techniques for STLF are explained in the time series classical approach. At the same time, STLF is subdivided into two different techniques. One is the time-series approach and the other one is artificial intelligence; some other hybrid approaches are also introduced in this regard, but they are not in the scope of this dissertation.

2.1 Time Series techniques for STLF

Early research started on short-term load forecasting was the time series, but with time when the nonlinear loads increased on the consumer side, these methods become obsolete and not that much accurate. On the other hand, artificial intelligence methods seem to show remarkable performance in short-term load forecasting.
The following methods used in time series:

- Multiple regression.
- Exponential smoothing.
- Iterative reweighted least-squares.
- Adaptive load forecasting.
- Stochastic time series.
- Autoregressive model (AR).
- Autoregressive moving average (ARMA) model.
- Autoregressive integrated moving average.

2.1.1 Multiple Regression

Multiple regression is a statistical method for predicting dependent variables as a linear combination of one or more variables called the independent variable. In this method, multiple variables were considered like weather, season, or day and the results were estimated or predicted by using a definite amount of historical data.

Many examples of multiple regression methods, such as [17], used this model in his research and showed quite satisfying results by implementing multiple regression in STLF. Whereas [6] has performed a literature review on short-term load forecasting models.

$$y_t = v_t a_t + \varepsilon_t \tag{2.1}$$

Where

t= sampling time

y_t = measured system total load

v_t= vector of variable that depends on the day (workday/holiday)

a_t= transposed vector of regression co-efficient

ε_t= model error at time t

In the above equation, the maximum polynomial degree is five that can be chosen for different variables that affect the forecast. It is seen previously that the linear dependency gives some time very accurate results.

For short-term load forecasting, previously [18] has implemented STLF for 24 hours and compared five different models, and discussed their implementation and results on the same database. Their experiments show very interesting results, e.g., for the hottest day in summer, the transfer function approach gives the best results; conversely, at the same time, the transfer function approach gives the second-worst result on the coldest day of winter.

A newly suggested model works like it will calculate the peak forecast of each day and then use this data of peak forecast to generate starting hour-based forecast [19]. From these calculations, they get the data of recent forecast error and exponential errors as variables in regression models, which will in return, produce altered peak forecast. To forecast the load more precisely, there exists a transformation technique that will take all changes in the data set of a load during the year, like seasonal change and annual growth of the load during the year [20].

Takeshi et al. worked to find out the peak load prediction and forecast this peak with transformation function or techniques for upcoming days using data from previous data sets. They suggested regression-based daily peak load forecasting methods with a transformation technique [20]. Their experiments performed a forecast considering dynamic elements like seasonal behavior, etc. and formulated a transformational technique. Different variables are calculated from past data sets and then transformed into new year data sets, which are used to calculate the new forecasts in this transformation technique.

Takeshi et al. extended the research on peak load forecasting. He presented a trends mechanism to estimate the peak load. For this research, he showed that peak load is much dependent on the weather conditions and humidity, so when the weather changed, there is a transition in the load behavior to so he suggested two trends techniques, one is trend cancelation, and the second one is trend estimation [21]. These techniques estimate and cancel the past load data sets with the new data sets and make the prediction more accurate.

Hyde et al. presented a load forecasting model based upon weather conditions, economic growth, and variable, which are dependent and independent on weather, like how the load will vary as weather conditions change; this model was designed to forecast the electricity demand of Ireland. They were focusing on the day ahead forecast, but this system was also applicable to 7 days[22] as they were using the weather data, so they use regression analysis for both weather and load. They used the past data available for load and weather and applied linear regression to estimate the future forecast load.

Alfares et al. presented regression-based each day maximum load demand for a whole year, including holidays. In order to predict load more accurately, they have to consider the following variable factors like weather. In the winter season, load behavior is not the same as in summers, and also, during the translation of seasons like spring and fall, the transformation technique is used. On holidays the effective load is subtracted from the actual load to calculate the load on holidays more precisely[23].

2.1.2 Exponential smoothing

Exponential smoothing works in a way, for instance, by using a previous load data and by modelling the future load forecast in which the model uses decreasing weights for past data set values when observation becomes old.

Exponential smoothing is categorized as one of the classical techniques used in predicting load. In this technique, the past data is being collected and then used as a transfer function to forecast the load of upcoming days. This exponential smoothing model used by[18], represented the following equation shown below. The electrical load at time "t" and y(t) is shown, and he uses the fitting function as follows,

$$y(t) = B(t)^T . F(t) + \varepsilon(t) \tag{2.2}$$

Where

$f(t)$ shows fitting function in the process

$\beta(t)$ represents a coefficient vector

$\varepsilon(t)$ shows the white noise

T is the transpose

The Winter's method was categorized as one of the exponential smoothing methods. It uses the different dependent variables of the season to produce a forecast. It can analyze time-series periodically easily. As in later analysis by Barakat et al. [24] It is very difficult to predict the load demand in a specific area that is growing very fast. So they suggested a model Short term load forecast in a fast-developing utility with inheriting dynamic load characteristics.

A new approach was developed in response to trend estimation and trend cancelation by Infield et al. The concept behind this approach was based on a specific fixed interval or optimal smoothing techniques[25]. Their result was very promising and showed more accuracy.

2.1.3 Iterative reweighted least-square

A proposed method using iteratively reweighted least-square to identify the model order and other related variables. In this model, they use an operator that can control only one variable at a time. Initially, this operator describes the initial point. This method uses autocorrelation and one-sided autocorrelation (partial autocorrelation) functions of previous electrical loads data and their residuals to devise a model for the load dynamics [17].

Consider a parameter prediction problem involving the linear calculation equation.

$$y = xB + \varepsilon \tag{2.3}$$

Where

y represent nx1 number of observations

x is nxp vector, matrix of known coefficient (Past data set)

B is the px1 vector of unknown variables or forecasted variables

ε is the matrix of random errors

It was concluded that the outcomes would be more accurate when an error doesn't lie in Gaussian surfaces and the forecasted variables must be obtained by the iterative models [17][6] consists of random error. They concluded that results would be more accurate when errors are not Gaussian B, obtained by an iterative method.

2.1.4 Adaptive Load forecasting

This technique of adaptive load forecasting means that forecasting variables can be updated or changed automatically as per changes in the load condition no need to add them or update them manually. Adaptive load forecasting can be used as an online software at the main distribution site, where it can control the variation in load and automatically update the parameters.

In regression analysis based on Kalman filter theory. This theory works like; it makes a series of measurements between specific intervals of time and some inaccuracies and produces an estimate of unknown variables that are more correct than other methods that use single values for predictions. In this theory, they experiment with all the data available to make correct forecasts like weather and load details; it will not work if, used the most recent data available for weather or load. This type of operation allows to switch between the adaptive and multiple regression analysis.

A solution presented to predict the nonlinear functional relationship between the load by Q, C. Lu et al. tried to use a Hammerstein function nonlinear relationship between temperature and load. An adaptive Hammerstein model with an orthogonal escalator structure and a lattice structure for collaborative processes is developed for short-term load forecasting from one hour to several hours in the future [26]. The benefit of this process was that they didn't use any matrix, so all complex calculation is being avoided.

The adaptive Hammerstein model was further studied, and improvements are made by Crawford et al. the main achievement in the upgraded Adaptive Hammerstein model was that its forecasting time was increased up to five days. The detailed implementation, forecasting procedure, and enhancements of this upgraded version Adaptive Hammerstein model are described[27].

One more enhancement solution was presented for adaptive short-term load forecasting, and that is named composite modelling. In this modelling technique, the load is sub-divided into three categories first one is nominal load and the second one is a type of load, and the final one is residual load. Kalman filter is being used as a parameter in the nominal part of the load. Other parameters of the load are adapted by the exponential weight recursive least square method [28]. The type load is basically for weekends and is being updated by the exponential smoothing method. Residual load is estimated by an AR model.

McDonald et al. presented a new short-term load forecasting technique for the distribution side of the power generation system [29]. In their research, they show how real the weather factor is adopted in short-term load forecasting. Their research implementation was carried out with several time series forecasting models like ARMA, WRSL, and VFF variable forgetting factor [23]. The model parameters were estimated as well as updated by WRLS. Finally, a software package was developed for the short-term load forecasting method for load distribution power utilities. Their online and offline testing has shown very good results with (MAE) Mean Absolute Error with less than 2% for 24-hour load forecast that was quite satisfactory and forecast with less than 2.5% error in the next 168 hours ahead[29].

Hyde et al. developed a model to forecast electrical load for the next 24-hour forecast; this model was also applicable for the next 5 to 10 days forecast. This model has been developed that consist of a rule that is dependent on weather-sensitive load component. They used the linear regression analysis of previous load and weather data to identify the load model [22]. They estimated the weather-sensitive component using the regression analysis. Their result shows a successful operation over time. They presented an automated load forecasting system that was with an additional feature of the design. Their algorithm was automated employed in both the identification and treatment of all those factors affecting load forecast directly or indirectly. In adaptive forecasting, forecasting is adaptive that model parameter is automatically corrected to keep track of changing load conditions.

2.1.5 Stochastic time series

It is obvious from previous research that, it is very difficult to predict energy requirement and type of load in very fast developing areas by using any one of time series method. But still these methods are helpful, by using time series to predict the electrical load of an area for short term.

First model suggested by using time series approach and the model was devised which were based on the past data sets. They use the previous load data to predict short term load in future. These time series models were only used for short term load forecasting methods.

2.1.5.1 Autoregressive (AR) model

Autoregressive model is used to predict future values, if there is a linear relationship between past values and future values of a variable that is expected to be forecast. In this case if the load is linear combination of past load details then autoregressive model is used to forecast the future load, K. Liu et al. has performed experiment on three different techniques for STLF[30] and Observed the results of his experiments by simulations in which the AR model is given by this equation,

$$L_K = -\sum_{i}^{m}.\ \alpha_{ik}.L_{k_{-i}} + wk \qquad 2.4$$

Where,

L_K represent the predicted load at time k

w_k shows the random load disturbance

a_i is the unknown coefficient from $i = 1 \dots m$

This equation is autoregressive model of order m. The unknown coefficient in above equation can be tuned online using well known least mean square(LMS) algorithm of [17].

2.1.5.2 Autoregressive moving average (ARMA) model

In Autoregressive moving average model the present values of time series $y(t)$ expressed in a linear relationship with the same time series values in the preceding periods $y(t-1), y(t-2) \dots y(t-n)$ and mean while considering the white noise of previous time slot as well, that can be represented by $a(t) + a(t-1) \dots a(t-n)$ [6] . In this scenario the ARMA model of order (p, q) can be written as,

$$y(t) = \emptyset_1 y(t-1) + \cdots + \emptyset_p\, y(t-p) + a(t) - \theta_1\, a(t-1) - \cdots - \theta_q\, a(t-q) \qquad 2.5$$

The parameter identification for the general ARMA model can be done by recursive scheme or using a maximum likelihood approach[6], which is basically a non-linear regression algorithm. Fan and McDonald used the ARMA model to represent the residual load of the power distributed system and WRLS (weighted recursive least square) to upgrade the variables of their adaptive ARMA model[29].

2.1.5.3 Autoregressive integrated moving-average (ARIMA) model

The Autoregressive integrated moving-average (ARIMA) model is used to predict the future load by examining the difference between the values in time series rather than the difference between the actual values. This method is mostly use for the conversion of dynamic process to stationary one and this transformation is performed by difference method. The series difference in AR autoregressive and moving average is done by inputs (p, q) but for ARIMA it is (p, d, q) than the model can be written as

$$\phi\ (B)\ \nabla \alpha y\ (t)\ =\ \theta\ (B)\ a\ (t) \qquad 2.6$$

The ARIMA method is used by Barakat et al. to analysed and modelled the past dataset values by ARIMA, which in turn used to future forecast of load demand with seasonal variation [24].

Juberias et al. worked on a new ARIMA model for STLF [31]. The model is based on time Series analysis methodology and includes the action of meteorology as an explanatory variable by using real time load forecasting ARIMA. Information about the influence of meteorology on hourly electrical load is given to the model as an explanatory variable by using the daily electrical load forecast. This is chosen as the explanatory variable because it can be considered as the most efficient and simple way to estimate the influence of meteorology in every hour of the day. The use of the model in a real time control system is also explained in the work presented by [31].

2.2 Artificial Intelligence Techniques for STLF

Research on artificial intelligence was started in last few years and due to its marvellous success in this area of short-term load forecasting, this topic is getting hot favourite for new researchers. Artificial intelligence has solved many real-world engineering problems in which our main focus is towards STLF.

There are various AI models for STLF, some of them are as follows,

- Fuzzy Logic

- Neural Network
- Knowledge Based Expert System

2.2.1 Fuzzy Logic

Fuzzy logic is a type of computing that predict the results, that depends upon degree of truth unlike Boolean logic 0 and 1 on which our most of computers works. Fuzzy logic is also used in energy forecasting models. Through this logic one can identify any unknown dynamic system, i.e. electrical load in this case, to some arbitrary accuracy. Infact there are many periodic changes in weekly load trends during weekdays, weekend, month, seasons, and so on. The fuzzy logic proved that it has great capabilities in drawing similarities from huge data[6]. So if there are enough historical input-output data pairs is available then the similarities existing in load trends are able to be identified. K. LIU et al. performed a comparison of three short term load forecasting techniques and concluded that fuzzy logic and neural network can be very good candidates for STLF[30]. They observed that fuzzy logic can forecast resemblance very precisely from a large amount of data. Resemblance in inputs $(L_i - L_0)$ of a source can easily be recognized from first order difference V_k and second order differences A_k which can be represent as;

$$V_k = \frac{L_k - L_{k-1}}{T} \qquad 2.7$$

$$A_k = \frac{V_k - V_{k-1}}{T} \qquad 2.8$$

The proposed fuzzy logic-based forecast works in two stages that is training and online forecasting. In training stage, the metered historical load data are used to train 2m-inputs and 2n-outputs to produce load forecasted by fuzzy logic using the first order and second order difference. After if there are sufficient number of training then it is linked with the controller to predict the load change online. If the most probably matching pattern with the highest possibility is found, then an output pattern will be generating through a centroid defuzzifier[6].

A research based on fuzzy logic was made to minimize the model error as well as to understand the nonlinear behaviour of power system. Hiroyuki et al. proposed an optimal fuzzy inference method for short-term load forecasting. The proposed method constructs an optimal structure of the simplified fuzzy inference[32].

As all forecasting techniques works quite well on normal days but on weekends or on special holidays error is quite high. So a new hybrid research based on fuzzy logic and neural network was presented, the main focus in this research was to minimize the error that occur in load forecasting on start of weekend and at the end of weekend or public holidays etc. The Fuzzy-Neural Network (FNN) has been extensively tested on actual data obtained from a power system for 24-hour ahead prediction based on forecast weather information. Very impressive results, with an average error of 0.62% on weekdays, 0.83% on Saturdays and 1.17% on Sundays and public holidays have been obtained [33].

K. Padmakumari et al. proposed a hybrid fuzzy neural network technique. The main focus of this technique lies in its ability to reduce applicable computational time and accuracy with other forecasting techniques[34].

2.2.2 Neural Network

The studies of neural network begin from last five decades but their real applications were found in last few years and the research on this topic is still on going.

Neural network is very use full tool for forecasting / prediction. Neural nets contain simple elements operating in parallel, these elements took inspiration from the working mechanism of nervous system of humans. The artificial neural networks can be used and trained, to perform a particular task and then by adjusting their input weights values they achieve particular target (outputs) very accurately.

Neural network or artificial neural network have very large number of applications, because of their ability to absorb new techniques by training algorithm. There are multiple hidden layers in these networks, in each hidden layer there are many neurons and inputs are multiplied by their weights. And later they are added to thresh holds to form the inner product number called the net function. The main advantage of this model is that the main model does not require the type of load information as it is required in other models.

A lot of researchers proposed short term load forecasting technique using neural networks because this technique has proved to be the best for the prediction of a short-term load demand. This method involves several variables that has to be chosen by accuracy. These variables involve one variable to determine the weather condition, one is use to determine the hour of the day, one is used to determine the season of the year and one is used to determine the previous load details. To enhance the ability of artificial neural network (ANN) to learn, they can easily manipulate the previous load data and predict the future load demand for short term load forecasting. ANN are quite able to learn the relationship among the past, present and future weather variable's and load combining both time series and regression approaches[35].

In load forecasting the term load modelling is very important for the demand management department therefore they should be aware of the class of the consumers, business clients, commercial or residential clients. A modelling approach is being presented to model load on power distribution side, and forecasting the load[36].

Nonlinear load is always a challenge to the load forecasting system in this regard a new comparison of two nonlinear STLF is made by Philippe et al. [37]. The two nonlinear techniques they included in their research are NN and Gaussian Process (GP) regression models. While the Bayesian approach to NN modelling offers significant advantages over the classical NN learning methods, it is concluded that the use of GP regression models will improve the performances of the forecasting.

A new proposed hybrid technique to improve the short-term load forecasting using neural network and by introducing several innovative features which predicts hourly load change, these features are created by historical load data and enhance performance of suggested algorithm to predict hour ahead electrical demand. They introduce a feature selection algorithm RReliefF used for selecting more (relevant) dependent feature on STLF. They employed multi-layer perceptron neural network because of their ease in integration, fault tolerance and self-organization [38].

Zhanle Wang et al. proposes a full wavelet neural network approach to forecast electrical load for short term. This method encompasses the full wavelet packet and neural network transform. The full wavelet transform technique further divides the electrical load profile and separates different features that alter load demands with distinguishable frequencies that we can use to teach the neural network to perform load forecasting for short term.

This proposed algorithm reduces the mean absolute percentage error by 20% if we compare it with other neural network techniques[39].

Today Neural networks are performing very well in solving the problems that are hard though to solve empirically.

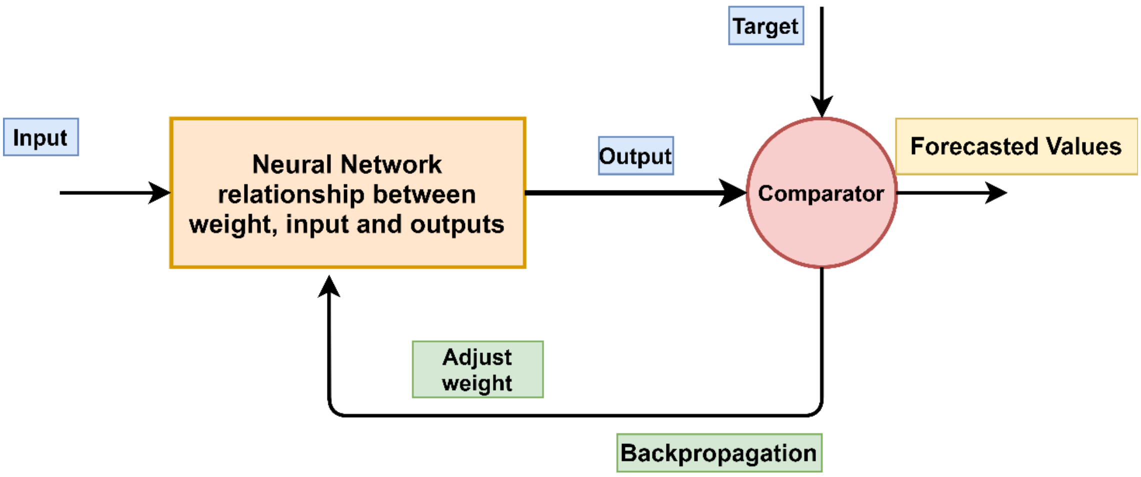

Figure 2. General work flow diagram of Artificial Neural Network[40]

Figure 2, shows the general work flow of the neural network architecture. This is how the inputs are subjected to the network and after weights are being multiplied with inputs, then the results passed through transfer function. Finally, these values are compared with target values, if the predicted values are not close to target values, then through backpropagation mechanism, weights of inputs are being adjusted to make the forecasting values closer to the provided targets.

Simple as well as very complex functions can be performed through neural nets. List of neural network applications includes Business Applications, Aerospace, Automotive, Banking, Credit Card and Activity Checking, Defence, Electronics, Entertainment, Financial, Industrial, Insurance, Manufacturing, Medical, Oil and Gas, Robotics, Speech like speech recognition, Securities analysis like in stock trading advisory system, telecommunications etc.

If linear functional relation is expected to perform between the input and expected targets outputs, then it is recommended to use the linear transfer function that is purlin. If a nonlinear transfer function used for a linear computation, then nonlinear transfer function will not perform well and the forecasting error will be high.

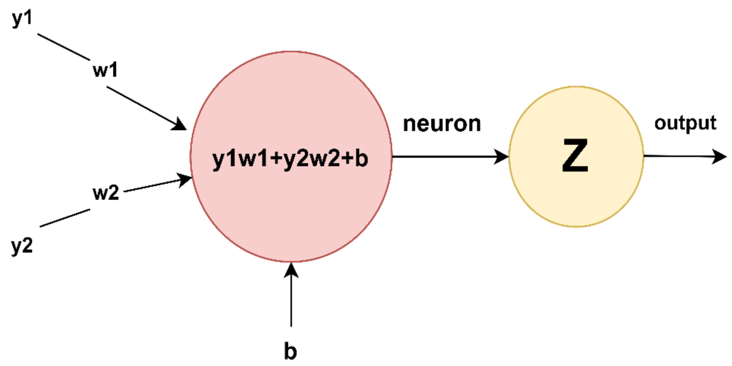

Figure 3. Relation between input, output, weights and bias of ANN

The figure above shows that input vectors from $y1, y2$ are multiply with their weights and then added. The neuron has bias 'b' which is added with weights then generates the net input neuron. which is then finally injected into the transfer function Z, whereas the output depends upon type of transfer function is used in the network.

This relation between weights, inputs and biases can be represented by equation given below,

$$n = W * y + b \qquad 2.9$$

For large number of inputs, the equation cab ne written as,

$$\boldsymbol{n} = w1{,}1y1 \ + \ w1{,}2y2 \ + \cdots + \ w1, ryr \ + \ b \qquad 2.10$$

$$\text{Weight Matrix} = \begin{bmatrix} W1{,}1 & W1{,}2 \ldots & W1, R \\ W2{,}1 & W2{,}2 \ldots & W2, R \\ WS, 1 & WS, 2 \ldots & Ws, R \end{bmatrix}$$

2.2.2.1 Application of Neural Networks

- Aerospace Application: Neural networks are used in air craft path trajectory simulation, also used in autopilot mode, air craft fault location detection.

- Automotive Application: In the advanced automobiles neural network are used automatic path guidance like parking and self-driving.
- Banking Application: Very useful in Credit application and evaluation.
- Credit Card Activity Checking: they can be used to check the spot of unusual credit card activity, which might be possibly result in credit card loss.
- Defence Systems: Neural networks are also used in defence system and Multiple input multiple output radar system, Discrimination of different objects. Face recognition, noise suppression etc.
- Electronics: ANN are used in chip failure analysis as well as nonlinear modelling of different electronic systems.
- Entertainment: Neural Networks are used in animation, editing etc.
- Financial: They are used in currency price forecasting, as well as financial analysis.
- Industrial: Neural networks play a vital role in forecasting electrical load demand for industries, they are now also used to predict gases coming out from chimneys of different furnaces, so they saved a lot of money because now there is no need to buy very expensive gases detective devices.
- Insurance: Neural networks are used in policy evaluation.
- Manufacturing: They are used to test the quality of welding analysis, chemical production analysis, and cement industry to forecast the load demand.
- Medical: Neural Nets are used to analyse the breast cancer cells; they are also used in ECG analysis.
- Oil and Gas: They are used to forecast and analyse the reserves of oil and gas, as well as exploration of oil resources.

- Robotics: They are used in robotics as vision system, they can differentiate the path with less obstacles and more obstacles and then robot follow the path with less obstacles.
- Speech: Neural networks are used in speech recognition.
- Securities: Neural networks are used in facial recognition.
- Transportation: Neural networks are used in truck brake analysis system

2.2.2.2 Transfer Functions

- Hard Limit Transfer Function
- Purelin Transfer Function
- Log Sigmoid Transfer Function

Hard Limit Transfer Function: The hard limit transfer function has ability to subdivide the input. The output will be zero, if input is less than zero and the output will be one, if the input is greater than zero or equal to zero, as shown in figure 4.

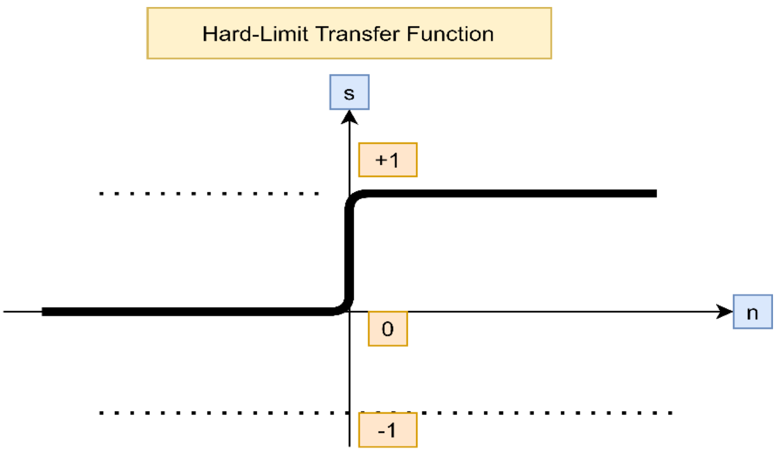

Figure 4. Hard limit Transfer Function graphical representation[40]

Purelin Transfer Function: For a linear function it is recommended to use the linear transfer function that is known as Purelin. Purelin is always used for linear networks. The main drawback of linear network is that they can only solve linear problem. In other words, neural networks with linear relationship are called without the brain stuff. As they have no back-propagation mechanism so they cannot update their outputs with respect to target (inputs).

If there is not a linear relationship between inputs and outputs, then backpropagation is good option to consider. Purelin transfer function is shown in figure 5.

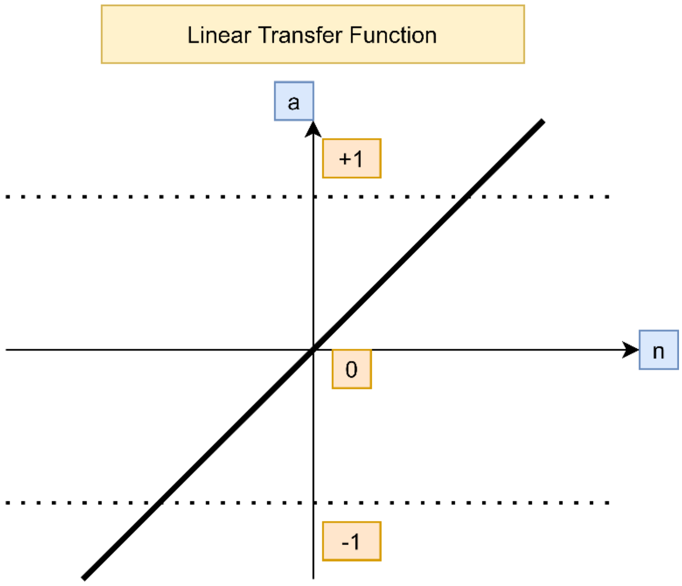

Figure 5. Purelin Transfer function graphical representation[40]

Log Sigmoid Transfer Function:

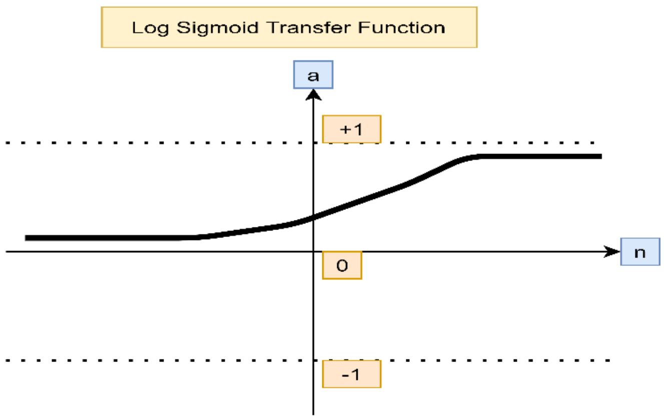

Figure 6. Log sigmoid transfer function graphical representation[40]

In a Neural Network, a log sigmoid transfer function is used to limit the response at each node of the network, the output of a node being the sigmoid of the weighted sum of the inputs. The sigmoid transfer function is mathematically a convenient way to limit the output and to adapt the weights to provide a learning function (backpropagation).

Simple Neuron Structure

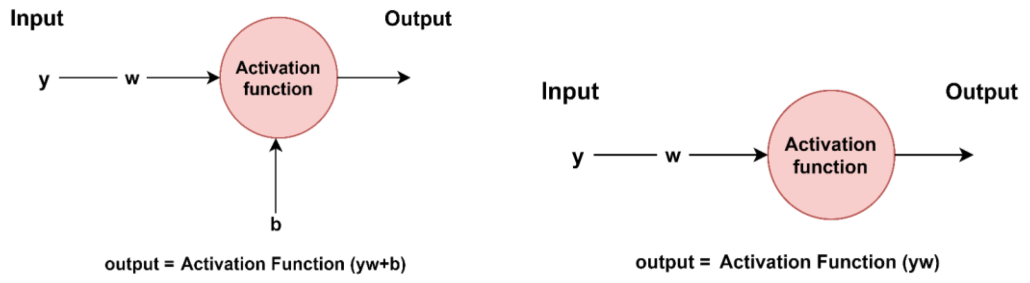

Figure 7. Single neuron with bias and without bias

In figure 7, on the right side, there is no bias. The input p is multiplied by weight and then added to transfer function and gives output neuron, but the other figure on the left side there is a bias that is being added to input after being multiplied by weight, this bias will shift the output toward left by amount b.

Actually, this architecture is the representation of network, like how many neurons does the network have or how many layers are there in the artificial neural network, what is the transfer function of network, and finally how the network layers are connected to each other.

Architecture of neural network vary from network to network and it depends upon the type of problem that is expected to solve. Number of neurons in output layer are always fixed while number of neurons in other layer of neural networks depends upon the problem trouble-shooter.

Neural network that uses the biases can show the link between input and output more precisely as compared to the neural network without biases.

2.2.2.3 Layers of Neurons

In this figure 8, all the inputs are connected to all the neurons with different weights. Each product of input vector ‘z’ and weights after addition of bias becomes arguments of transfer function and then after passing through transfer function generates output.

If the input vector is categorized, then there will be two types of input vectors one is concurrent input vectors and second are sequential input vectors. Concurrent input vectors are those in which input order is not mandatory, if network is running in parallel, then it is easy to assign one input to each of the network. The sequential input vectors, the order of input coming to network is very important. Multiple layers of neurons are shown in figure 8.

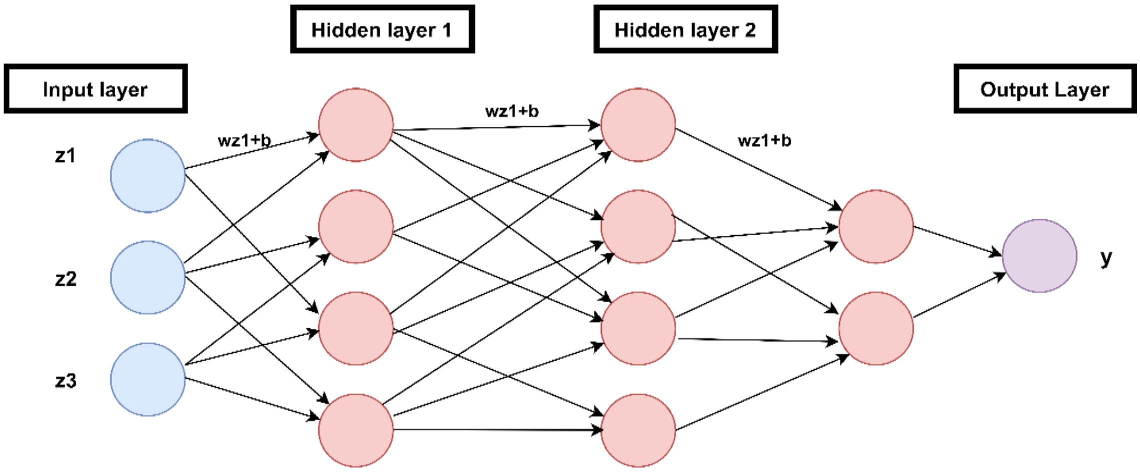

Figure 8. Multiple layers of neurons, input layers, output layers and hidden layers

2.2.2.4 Training of Neuron

The training function of neural network contains a series of calculations, in each cycle the function pass through specific sequence of inputs and calculates the outputs and errors, finally adjust outputs for each sequence of input presented.

Training does not make sure the network, that is being created does its job. Consequently, it will be a good practice to check the output with respect to target inputs, and compare those achieved results with targets and conclude that, results are successful or not with respect to mean absolute percentage error. If there is a big difference between the targeted input and network input, then retrain the network many times until unless the expected results are achieved.

Training Styles of neural nets

- Incremental Training
- Batch Training

Incremental Training: Incremental training is used in both cases static networks as well as dynamic network. Moreover, it is commonly used with dynamic networks.

Batch Training: This type of training can be done by updating weights and biases after all the inputs and targets are applied in network to train. Batch training can also be applied in both static as well as dynamic networks.

Single neuron has not enough power to analyse or forecast anything, so when multiple neurons combine that have really good computational or forecasting power.

Steps of Training

There are sequences of training steps that need to be followed while training the neural network, these steps are as follow,

- Training data must be assembled.
- Creating the object of a network.
- Train the network.
- Network response simulated with new inputs.

2.2.2.5 Learning Rules

Learning rules are actually the training algorithms of updating weight and biases of a network. Different learning rule are applied to the designed network to perform specific tasks. To be more precise, it is the method in which the network adjusts weights and biases of inputs to make the outputs closer to target inputs. There are two types of learning schemes, Supervised learning

Unsupervised learning

Supervised learning target outputs are provided, as the inputs are already applied to the network so the network will compare outputs with the target inputs, and adjust its weights and biases, and generate more precise results.

Unsupervised learning, the weights and biases are altered as per inputs values because there is no need to provide any target input in unsupervised learning, to cut it short there is no target input available in this scheme of learning.

Perceptron Learning Rules

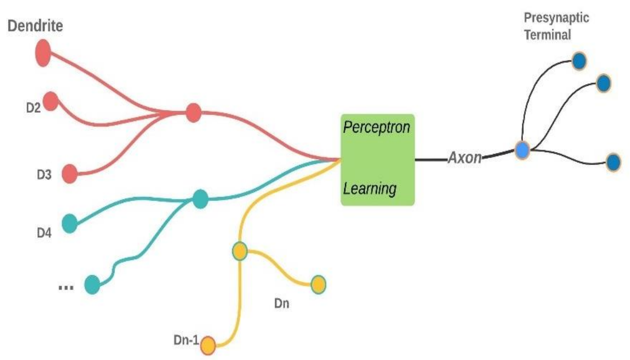

Figure 9. Perceptron learning mechanism

The above figure shows that, Dendrites with multiple inputs and helps to carry the impulse response from different parts of the body, these responses are carried through cells, which are neurons. Each neuron is made up of long thread, even thinner than human hair, this thread is called axon. The information passes through axons and gives output signal on presynaptic terminal.

The perceptron can be trained with their input and targets. This relationship can be represented as,

$$D1t1, D2t2, \dots, Dntn \tag{2.11}$$

Where 'D' is an input to the network and 't' is the corresponding target (output) of the network. The main task is to minimize the error that is actually obtained by subtracting neuron response ‘a’ from target, ‘$t - a$’. If this error will be bigger than its mean that the network is not trained well, the network should retrain again, until unless minimum error is achieved.

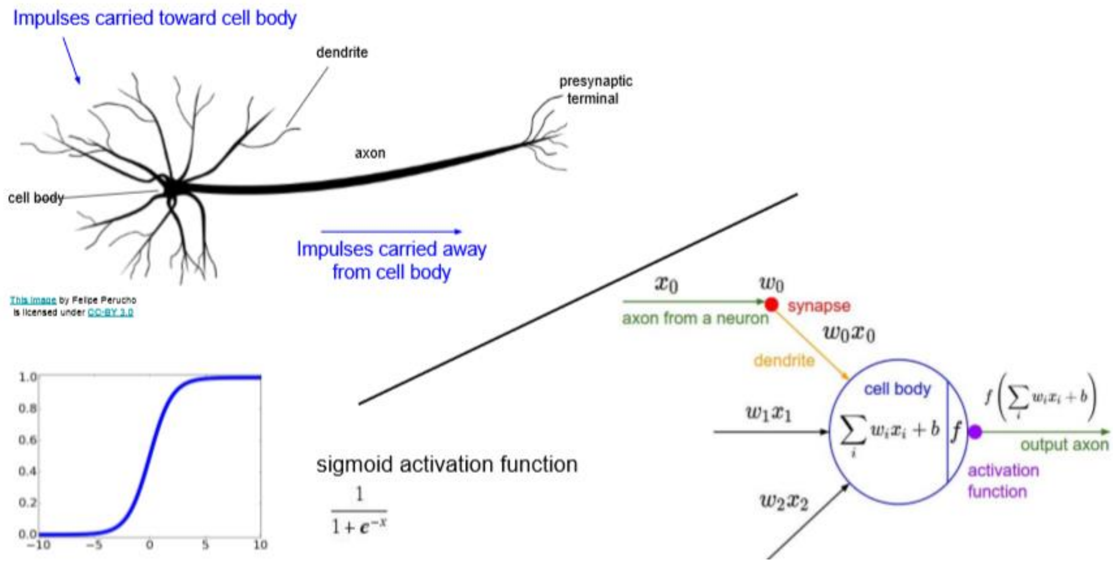

Figure 10. weights, biases and inputs computational relationships[40]

The perceptron learns through learnp and calculates the desire changes to weight, biases for the given input and the error ‘e’. The target always contains values between 0 and 1 for hardlim transfer function.

If the biases were not introduced in perceptron learning, then the perceptron works to find out the solution by only altering the weights of the inputs vectors and classify them as zero or one. The points toward the inputs are classified as ‘1’ and those points which are away from the input are classified as ‘0’.

Actually, there are three conditions which results, when an input is applied to neuron.

Case 1. If input is applied to a network and output of the network is correct, like exactly expected response is obtained from the neuron. ‘a’ represent the neuron output and ‘t’ is target (a=t, and error = t-a =0), in this scenario the weights will not be changed by perceptron. This is the ideal condition when the error is exactly zero, that’s not possible in real world problems. The error can be minimized but never completely eliminated from the forecasting system.

Case 2. If neuron gives the output zero but it must be 1 as per target values then (a=0 and t=1, and error=t-a=1) in this scenario the input vector ‘p’ will be added to the weight vector ‘w’, this will make the weight vector closer to the input vector, so that there will be more chances in next prediction near future that input vector will be classified as 1.

Case 3. If the neuron output is 1 and it must be 0 as per targets (a=1, t=0 and e=0-1=-1), hence it means that the input vector ‘p’ must be subtracted from the weight vector ‘w’ this will result in moving the point further away from input vector, accordingly it will be classified as 0 in future.

2.2.2.6 Feedforward Network

Feed forward network mostly contains two or more layers of neurons, hidden layer and output layer. Hidden layer is composed of sigmoid neuron and the output layer is made up of linear neuron. These multiple layers of neuron with different transfer function on input side and output side makes the network more understandable the linear as well as nonlinear relationship between inputs and outputs. If it is needed to limit out output values between zero and one, then sigmoid function is best to choose on the output layer.

2.2.2.7 Backpropagation Algorithm

Backpropagation was made by considering the generalization of Widrow-Hoff learning rule to network with multiple layers as well as nonlinear distinguishable transfer function. Inputs and the parallel targets input vectors are used by the network to train itself, and approximate the function that will generate the output much closer to the target input.

Benefit of backpropagation network is that it generates much satisfactory result for the inputs that are never subjected to network before, just by adjusting the weights of inputs as per required outputs.

Backpropagation algorithm updates the weight and biases in specific part of a network, which are causing the degradation in performance of network.

One complete cycle of this process can be represented as,

$$x_{k+1} = x_k - a_k g_k \qquad 2.12$$

Where,

x_k represent the weight as well as bias of network

g_k represent the current gradient

a_k is the learning scale

This algorithm can be implemented by two ways: incremental training mode and batch training mode. In incremental mode gradient is calculated and weights are updated after each and every input that is subjected to the network. In batch mode all inputs are applied to the network first then weights are being updated.

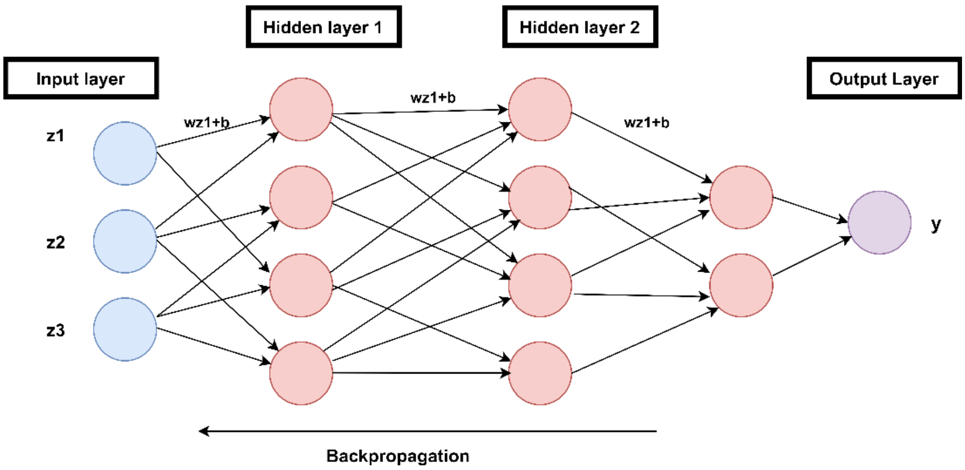

Figure 11. Backpropagation working principal of ANN

2.2.2.8 Quasi-Newton Algorithms

Newton method is utilized for fast improvements in learning speed of a network[40] . The newton method can be shown by following basic step,

$$x_{k+1} = x_k + A_k^{-1} g_k \qquad 2.13$$

Where,

A_k represent the Hessian matrix of second order

x_k represent the weight as well as bias of network

Benefit of Newton method is that its convergence is much faster than those of gradient methods, on the other hand, it is very difficult and expensive to do same calculations by using Hessian matrix for feed forward neural network.

There are several algorithms that are based on Newton's method in which there is no need to calculate the second derivatives, so all those methods are known as quasi-Newton methods. The quasi-Newton method is most impactful with respect to Broyden, Fletcher, GoldFarb, and Shanno (BFGHS) updates[40].

2.2.2.9 Levenberg-Marquardt

Levenberg-Marquardt method was developed on same functionality like quasi-Newton methods, to achieve second order speed of training the network without calculating the Hessian matrix[40]. In specific feed forward network the Hessian matrix can be approximated as,

$$H=J^T . J^1 \qquad 2.14$$

In the above equation 'J' represent the Jacobian matrix which is made up of first differential of network errors with reference to weight and biases of the network. This Jacobian matrix is easily calculated by using some backpropagation techniques. Calculating Jacobian matrix is far less complicated as compared to calculating the Hessian matrix [40].

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \qquad 2.15$$

Where,

x_k represent the weight as well as bias of a network

J represent the Jacobian matrix

μ represent the scalar quantity

e represent the error

μ is a scalar quantity and when its value is zero the remaining equation represents the Newton method with approximate Hessian matrix.

When the value of scalar μ is large the gradient descent form, it has small step size. The point of iteration where the error is minimum the Newton method becomes more accurate and faster. Therefore, our target is to switch toward Newton method as soon as possible.

On other hand the μ is decrease after each complete iteration which intern mean the reduction in execution time of the process. The value of μ is increased only in the start of training, which will increase the execution time but after words it will start dropping out until the process is successfully completed.

2.2.3 Knowledge based expert system

As the name suggest this is a technique in which computer is programmed that has ability to explain, extrapolate the new and previous information that is available to it. As this all is happened due to continuous development and innovation in AI, artificial intelligence.

To make this model, the previous load details are being extracted and then new rules are programmed for change in load conditions, weather condition, peak hour or base hour, changed in some forced conditions that directly effects the use of electricity, which intern effect the forecasting.

2.3 Conclusion of all state of art methods

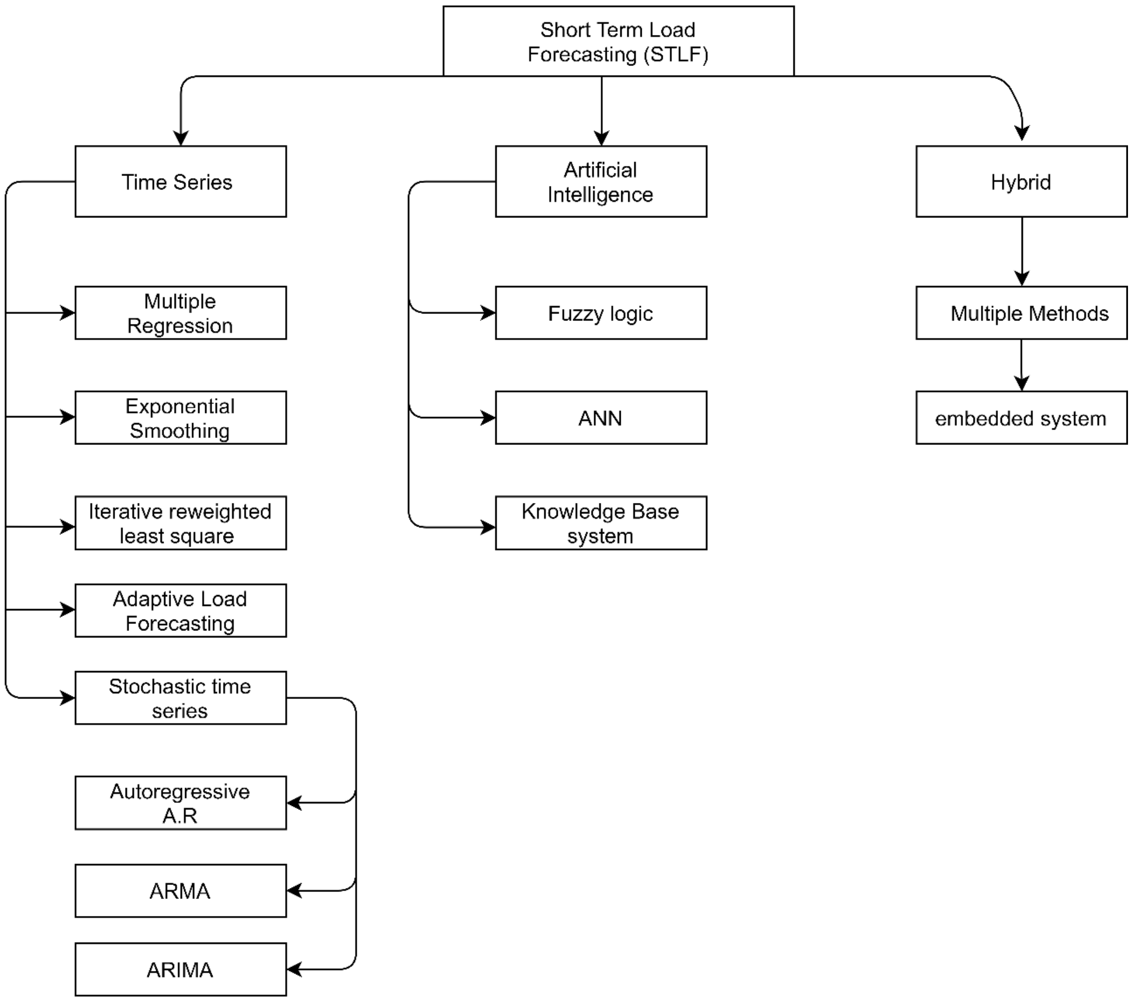

Figure 12.All STLF techniques reviewed in this dissertation

3 Methods

In this chapter, the methodology is discussed, about all different scenarios, using different conditions performed on short-term load forecasting through an artificial neural networks. Many researchers used the quantitative and mixed methods approach to solve the problem of STLF. Therefore, in this research by utilizing surveys of previous short-term load forecasting techniques, and manipulating different recognizing patterns, various experiments are performed using a ANN tool in MATLAB2020b, and calculations are done on past and future data sets values.

3.1 Data set collection

Data is collected from different sources described below.

3.1.1 Weather data

Data set obtained from different sources, e.g., the temperature, pressure, humidity data is obtained from the website rp5.ru [15][41]. This website is designed, maintained, and updated by St. Petersburg, Russia, from 2004 till now. This company has a license in hydrometeorology [17]. This website provides all the historical weather data. In addition, this website also provides almost one week's forecasted values of weather and its related parameters. Therefore, the data is being downloaded, including all historical data from 1st January 2021 to 18th March 2021, for the first three experiments, and due to failing in achieving the expected accuracy, the fourth and final experiment is performed using the data set values, increased from three months, to one year and three months to accomplish more accurate results, the dataset from 1st January 2020 to 18th March 2021 are taken into account for experiments, the data is downloaded in Excel datasheet format from rp5.ru.

There were many weather variables (T, P, U, wind blow etc.) in that historical datasheet. However, in this research for STLF, the main consideration is given to the three most effective variables: temperature, pressure, and humidity. Previous research has shown that these three parameters are the most worthwhile for consideration in STLF.

3.1.2 Electrical consumption data in Estonia

The second data set used in this research work is the actual consumption of electrical load in Estonia, supplied and managed by NordPool [16][18]. All the data related to electricity consumption is available publicly on their website. NordPool is the top power market in Europe, especially in the Nordic states. They provide all electricity-related facilities to their customers.

3.1.3 Other Inputs

Other inputs used in these experimental scenarios are days, months, and hours of the day. In order to differentiate between the hours of a day, in this dissertation, numeric representation for each hour is used, e.g., The number of hours is represented from 0 to 23. In order to differentiate between hours for different days, e.g., days are represented from Monday to Sunday with numeric digits from 1 to 7 in this research. Accordingly, the neural network can easily understand which hour and which day data are being subjected to them as input and for which day they are expected to forecast.

Secondly, the neural network will not understand the difference between the months, and for example, for which month data is provided and for which month ANN are expected to predict; therefore, in this dissertation, to distinguish between different months of a year, months numbers also used as inputs parameter, therefore numeric digit '1' is assigned to January, '2' to February and '3' to March and so on. In this way, inputs will be much more distinguishable, and accuracy will be high.

3.2 Importance of data sets

Data sets are very important for the artificial neural network; the more data sets value are used, the more precisely ANN is expected to train, validate, and test the network. finally, forecasted accuracy improves when neural nets choose similar data from large datasets as shown in scenario 4 in section 4.4.

3.3 Holidays/Weekends

The electrical load profile for working days of the week is almost the same with minor fluctuations in electricity consumption. However, a sudden drop in load consumption occurs at the start of a holiday. Consequently, this drop greatly affects the forecast, and the mean absolute percentage error will drastically increase, which is not a desirable condition in load forecasting.

Secondly, during a transition from a holiday to a Monday, the error profile is again high because the ANN is taking the load consumption of holidays as a target to forecast the load consumption profile for the working day (Monday).

To overcome this problem, in this research, many experiments have been performed to improve the accuracy and concluded the results, which are mentioned in experiments section 4.2.

3.4 Experimental Tools

All experiments in this research are done on a built-in toolbox of artificial neural networks in MATLAB 2020b. MATLAB stepped in the artificial intelligence spring and introduced many tools for solving real-time problems; one great example is the neural network toolbox. The neural network toolbox is an exceptional platform for deep learning applications.

3.4.1 Data presentation

In Neural network GUI, data is presented in two sets.

1. Input's data sets
2. Target data sets

Input data sets

The data is presented in the form of a matrix. Each matrix contains a different number of rows and columns. Rows act for data features meanwhile columns represent the number of samples.

Target data sets

The data in targets are the task or hint for neural networks to forecast the desire output values accordingly.

3.4.2 Outputs

After providing the input parameters and creating a network, different changes are applied in a network architecture as per experiment requirements like the number of Epochs and validations checks etc. Then the network is trained, validates and tests inputs, and generates forecasted outputs.

3.5 Overall Sketch of experiments

3.5.1 Different predictions

In this research, multiple forecasting approaches and different scenarios are used, like the scenario is through providing the previous day data e.g. actual consumption of load, temperature, pressure, hour and month of the year, as inputs to ANN, to forecast next day electrical consumption as outputs.

In some cases of forecasting, it depends upon different utility companies, sometimes previous day data is not available, so in that case, forecasting is being done with a different approach, for this problem, two days' previous data of actual consumption is used as an input to forecast the future consumption.

To extend this scenario of lacking the consumption data in utility company, in this research, experiments are performed with the load data available for the previous week (seven days' previous data set). The second experiment is based on forecasting using one-week previous consumption data. In order to enhance the accuracy of the forecasting model, three parameters are being considered which have a great impact on the forecasting, and those are temperature, pressure, and humidity, etc.

After failing to achieve the desired accuracy with the test data set of three months, the research criteria were extended by using the historical data set from one year and three months, to train ANN better. The outcomes of experiment 4, show very interesting results; artificial neural network accuracy increases with increased training data set values.

3.5.2 Relating it with a machine learning problem

All of the experiments performed in this research, like utilizing the previous day dataset, e.g., input data from previous day, input data from last two days and one-week previous data used to forecast the future load consumption. Thus it represents a linear dependency; like from previous data set values, ANN are expected to forecast the same data as an output for next week. Therefore, this shows a linear regression problem, as discussed in the second chapter. The second argument is that historical data is being presented to the ANN network as target data. Therefore, the network does not need to predict consumption independently; that is why the learning category lies in supervised learning.

3.5.3 Forecasting accuracy assessment

The forecasting accuracy of a model depends on how useful the output is or how close the predicted values are compared to the actual values. Secondly, in order to categorize the model performance, good, bad or excellent it does not just depend upon its accuracy of forecasting; it also depends on the dataset values that are provided to the forecasting model. If bad historical information is being subjected to a very good structured model, it will still generate a very poor forecast.

To analyze the accuracy for neural network forecast in this dissertation, the mean absolute percentage error (MAPE) is utilized, which is a statistical analysis to check the accuracy of prediction. MAPE always computes its percentage accuracy.

First of all, MAPE compute the difference between actual values and predicted values that are shown below,

$$F_{difference} = \frac{Actual\ consumption - Predicted\ consumption}{Actual\ consumption}$$

$$Absolute\ of\ error = |Actual\ consumption - Predicted\ consumption|$$

$$MAPE = \quad \frac{1}{n} . \sum_{z=0}^{n} . \ Absolute\ of\ error$$

3.6 Training, Validating and Testing

Forecasting is done through artificial neural network is divided into three steps as following,

1. Training

2. Validation

3. Testing

In the training phase neural network try to reach the optimal target values that are provided to the network, through the different coefficient, weights and biases that would help the neurons to achieve the desire outcome. Training does not make sure that our network is doing the specified task as per requirement. Training is further divided into two categories, incremental training and batch training. In incremental training inputs and biases are updated after each input applied to the network but in batch training inputs and biases are updated after all inputs are applied to the network. In order to train the network, it is suggested to provide a proper set of inputs, all inputs must be assembled. Selecting an input is not an easy task, consider all the variables very effectively that are directly affecting the forecast.

In validation phase of neural network, the input is applied to the neuron and it generates the output that is exactly same as like target output or desired output. In this case weights and biases will not be changed. In other scenario if neuron generates the output zero but as per specific target it must be one in this case bias is added to the neuron input and weight so in next cycle there will be more chances to make this output one. In this way adding weights and biases to neurons inputs neural network validated.

In testing phase, set of assembled inputs and desire target is applied to the network, then network generates an output values, actually those are the forecasted values. In order to check how accurate, the forecasting values, there are different techniques for error calculations. In this research work the forecasted error is calculated through mean absolute percentage error.

4 Results and Discussions

To select the best model, many experiments with different scenarios were performed by applying the different number of epochs to the network, different results obtained from the different epochs. It is not good to say that ANN performance depends upon the number of epochs. However, it is the right way to conclude that at which epoch the model is giving the accurate forecasting with minimum mean absolute percentage error. The mean absolute percentage error remains between 4.35% to 5.44% for 23^{rd} march 2021.

Experiments are done with an increasing number of epochs. At epoch number 50, the model shows the highest error percentage, and at 300 epochs, a model error was the lowest. Therefore, it is good to select 300 epochs at which ANN shows minimum mean absolute percentage error. The results are shown below in table1.

Table 1. shows different number of epochs and their MAPE result

Number of Epochs	Mean absolute percentage error
10	4.85 %
50	5.44 %
100	5.16 %
300	4.35 %
500	5.02 %

After finalizing the number of epochs, the inputs are subjected to the network and forecasted electrical consumption from 22^{nd} to 28^{th} march 2021.

4.1 Scenario I

In this scenario, previous day electrical consumption data is used as a target and other parameters like an hour, temperature, pressure, humidity, and months as inputs to the ANN model. Initial parameters of model are shown in table 2.

Table 2. Initial parameters setting of ANN according to scenario 1

Inputs of neural network	Number of hours of a day Temperature of the day Pressure of the day Month of the year Humidity Day of the week
Number of neuron in hidden layer	10
Activation function	tanh
Number of neurons in output layer	1
Activation function of output layer	Purelin (Linear)

After creating the ANN model with all the specifications, as shown above, the next step was setting the training parameter that are shown in table 3.

Table 3. Training, testing and validation parameters of ANN

Number of Epochs	300
Training data	15%
Validation of data	15%
Error estimator	MSE, MAPE
Testing Inputs	1 January-18 March 2021

After completing training, the network output forecasted results. Actual electrical consumption and forecasted electrical consumption is shown in table 4. '**A**' represents the actual consumption and '**F**' represents the forecasted output. The electrical load is measured in unit kilowatt-hours and represented by '**kwh'** in table 4.

Table 4. One week forecast from 22nd -28th march, with actual and forecasted values, using previous day electrical load values for experiment 1

Time	**Monday**		**Tuesday**		**Wednesday**		**Thursday**		**Friday**		**Saturday**		**Sunday**	
Hrs.	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A kwh**	**F kwh**	**A kwh**	**F kwh**	**A kwh**	**F kwh**	**A kwh**	**F kwh**
23	946	853.8	889	946.3	859	882.1	889	857.5	863	894.3	790	897	809	787.9
22	991	892.1	942	992.1	898	945.4	936	905.5	844	934.1	834	946.2	863	813.6
21	1063	941.2	1020	1061.1	977	1020.5	1014	975.7	836	1007.3	887	1011	919	865.9
20	1131	997.5	1098	1132.6	1050	1085.5	1076	1046.1	842	1070.2	947	1062	983	936.8
19	1180	1048.8	1138	1179.9	1097	1125.9	1121	1095.7	871	1118.1	995	1098	987	966.1
18	1190	1072.2	1157	1184.6	1126	1128.1	1138	1124.2	932	1102.9	1024	1106	928	967.1
17	1132	1047.2	1107	1149.7	1107	1088.1	1095	1110.1	1047	1109.6	960	1068.4	886	945.7
16	1093	986.3	1082	1097.7	1097	1057.6	1079	1093.5	1132	1112.3	930	1027.6	843	926.1
15	1057	933.5	1065	1056.2	1086	1060.5	1070	1092.9	1140	1072.6	913	1001	794	913.1
14	1040	908.5	1069	1037.5	1083	1074.6	1078	1088.6	1105	1062.3	916	981.4	775	904.3
13	1028	891.8	1070	1033.5	1082	1078.4	1084	1078.3	1078	1079.8	922	1007	794	898.2
12	1037	866.7	1083	1030.9	1067	1084.1	1102	1057.6	1040	1094.8	900	1032	829	892.6
11	1020	844.5	1079	1028.8	1018	1077.9	1101	1026.1	1032	1100	888	1040	884	889.7
10	1036	842.7	1104	1035.7	1001	981.1	1135	999.3	1007	1130.5	885	1052.1	924	885.2
9	1053	860.3	1127	1057.7	1014	1107.5	1148	1015.1	1003	1155.8	886	1090.1	926	881.1
8	1086	886.2	1153	1087.4	1029	1154.9	1153	1030.1	1001	1156.8	887	1140	893	878.9
7	1102	909.2	1168	1097.9	1055	1164.3	1130	1069.7	1028	1116.3	880	1132	815	873.9

6	1055	898.9	1105	1054.7	1018	1102.4	1033	1064.5	1055	1026.1	866	1047	757	870.3
5	955	848.5	1014	963.2	935	1014.3	923	998.2	1106	941.3	852	932	758	857.1
4	877	847.5	948	882.6	869	946.2	854	869.7	1098	865.2	856	865.4	757	844.8
3	849	855.1	919	845.1	835	917.2	822	836.2	1062	824.8	850	840.5	750	845.4
2	842	861.3	911	836.3	839	913.2	815	834.5	1011	810.6	849	836	745	855.8
1	832	865.1	914	837.1	842	913.4	817	850.8	938	812.5	854	844	742	863.6
0	840	868.2	923	838.3	850	922.7	826	850.1	897	815.7	871	863	750	849.9

The graphical representation of table 4, is shown in figure 13. It is can be seen that in the beginning of a week on 22nd March, on Monday the forecasting error is high, afterwards on working days there is not much difference between actual and forecasted consumption of load, but again on weekends, from hour 115 the forecasting error increases.

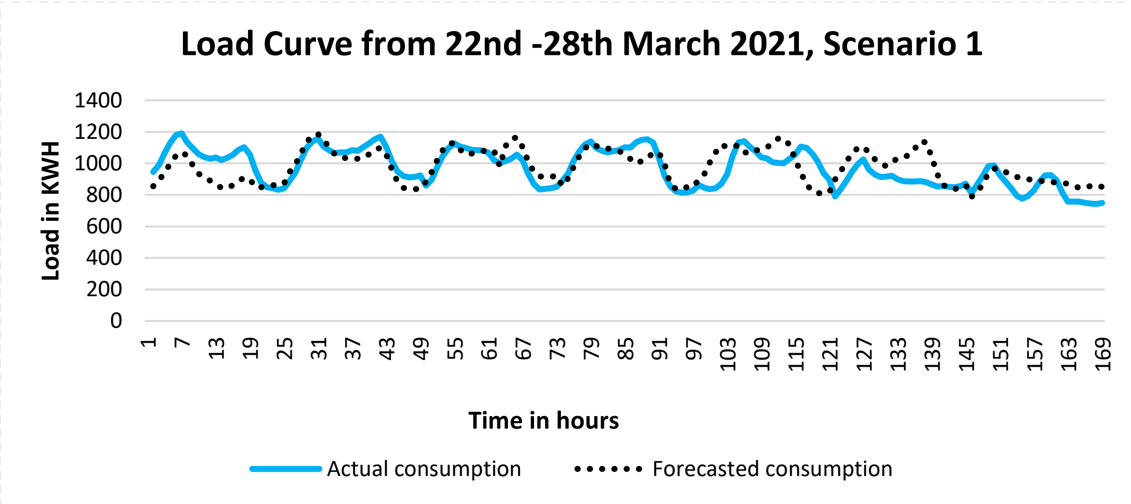

Figure 13. Graphical representation of scenario 1, between actual consumption and forecasted consumption

After forecasting for one week, finally computed the mean absolute percentage error for each day. The lowest MAPE calculated was 3.04%, so the best forecasting results were obtained for Thursday, 25th March, and the highest MAPE was calculated for Saturday, which was 12%. The MAPE calculation from 22nd March to 27th March is shown in table 5.

following formula used to calculates mean absolute percentage error,

$$MAPE = \frac{1}{n} * \sum_{i=1}^{n} . \frac{Actual\ value - Predicted\ Value}{Actual\ Value}$$

Where, n is the total number of time stamps, that are included in this forecast

Table 5. MAPE calculation for a week, using one-day previous data from 22nd march to 28th march for experiment 1

Days of week	**MAPE**
Monday	11.67 %
Tuesday	4.9 %
Wednesday	5.4 %
Thursday	3.04 %
Friday	4.2 %
Saturday	12 %
Sunday	9.8 %

In table 5, it is very obvious that on Monday, the forecasting error is quite high, as compared to the rest of the days, and it starts decreasing onward. The reason for high error on Monday is because of a commercial load of offices of different companies, schools, colleges, and universities are not operational on weekends, but on Monday, their load is added to consumption, therefore forecasting by using a previous day approach, ANN takes the weekends consumption (Sunday) as a target, which is not the same as working day (like Monday) so this is the main reason of high forecasting error on Monday.

ANN shows very satisfactory performance from Tuesday to Friday, but from Saturday, the error increased again.

4.2 Scenario II

In order to decrease the error, a different approach is applied for forecasting, like using target data for ANN from the previous two days, keeping in mind two reasons, one was to decrease a forecasting spike that occurs on Monday, and the second reason was that mostly the utility companies do not have actual consumption data for last two days, so these are two motivations for the second experiment. Input structure of the model is shown in table 6 and training parameters are shown in table 7.

Table 6. Initial setting of neural network, after experimenting and selecting the best suited model for experiment 2

Inputs of neural network	Number of hours of a day Temperature of the day Pressure of the day Month of the year Humidity Day of the week
Number of neurons in hidden layer	10
Activation function	tanh
Number of neurons in output layer	1
Activation function of output layer	Purelin (Linear)

Table 7. Training parameters of ANN used in experiment 2

Number of Epochs	300
Training data	15%
Validation of data	15%
Error estimator	MAPE
Testing Input	1 January-18 March 2021

Table 8, shows the hourly based forecast results for 24 hours of the day, from Monday 22nd march to Sunday 28th march. After getting, the forecasted results mean absolute percentage error was computed for each day starting from Monday to Sunday and MAPE results for experiment 2 are shown in table 9.

Table 8. One week forecast from 22-28 march, with actual and forecasted values, using 2 days' previous values of electrical load in experiment 2

Time	**Monday**		**Tuesday**		**Wednesday**		**Thursday**		**Friday**		**Saturday**		**Sunday**	
Hrs.	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F kwh**	**A Kwh**	**F Kwh**
23	946	1026.7	889	894.7	859	853.6	889	964.6	897	880.518	790	847.7	809	887.373
22	991	1053.3	942	942.7	898	887.6	936	990.3	938	941.98	834	897.9	863	945.5706
21	1063	1124.6	1020	1003.9	977	942.4	1014	1062.2	1011	1019.95	887	977.1	919	1017.138
20	1131	1179.8	1098	1073.2	1050	1004.4	1076	1120.4	1062	1097.9	947	1049.8	983	1071.642
19	1180	1223.8	1138	1120.3	1097	1050.6	1121	1181.9	1098	1138.03	995	1097	987	1131.26
18	1190	1247.6	1157	1152.7	1126	1058.7	1138	1189.3	1106	1107.05	1024	1109.4	928	1138.892
17	1132	1192.5	1107	1113.6	1107	1022.7	1095	1131.6	1055	1107	960	1094.1	886	1093.916
16	1093	1157.9	1082	1076.6	1097	984.6	1079	1092.2	1028	1081.86	930	1096.9	843	1079.127
15	1057	1123.8	1065	1045.6	1086	939.3	1070	1056.7	1001	1064.78	913	1086.1	794	1070.453
14	1040	1119.4	1069	1032.6	1083	914.3	1078	1035.6	1003	1068.80	916	1097.5	775	1078.538
13	1028	1118.1	1070	1020.6	1082	894.9	1084	1027.2	1007	1069.78	922	1082.01	794	1081.43
12	1037	1122.8	1083	1026.1	1067	912.2	1102	1021.2	1032	1083.49	900	1056.43	829	1097.866
11	1020	1174.7	1079	1024.4	1018	851.6	1101	1027.4	1040	1078.9	888	1005.36	884	1106.958
10	1036	1177.3	1104	1032.6	1001	835.9	1135	1034.7	1078	1104.04	885	1000.91	924	1127.979
9	1053	1226.8	1127	1038.9	1014	855.1	1148	1051.6	1105	1176.07	886	1014.01	926	1151.389
8	1086	1215.9	1153	1039.1	1029	875.8	1153	1085.	1140	1153.16	887	1028.90	893	1153.835
7	1102	1217.2	1168	1022.4	1055	901.7	1130	1100.1	1132	1128.8	880	1072.38	815	1127.555

6	1055	1130.1	1105	991.1	1018	909.4	1033	1053.3	1047	1098.16	866	1018.19	757	1048.553
5	955	1063.9	1014	973.3	935	880.7	923	954.6	932	1008.5	852	934.778	758	923.9023
4	877	951.8	948	966.1	869	858.7	854	877.1	871	948	856	867.694	757	836.5697
3	849	932.4	919	963.9	835	848.6	822	845.8	842	919	850	838.976	750	814.0987
2	842	909.2	911	966.3	839	853.6	815	836.3	836	911	849	836.201	745	811.9266
1	832	915.5	914	968.4	842	859.2	817	832.7	844	915.1	854	842.362	742	817.3646
0	840	919.4	923	987.8	850	869.4	826	833.3	863	923	871	853.478	750	824.6243

The graphical representation of table 8, is shown in figure 14, it can be seen that forecasting error on weekends from hour 115, is very high as compared to scenario 1. MAPE on Sunday rise to 20%, which is not acceptable in any good forecasting.

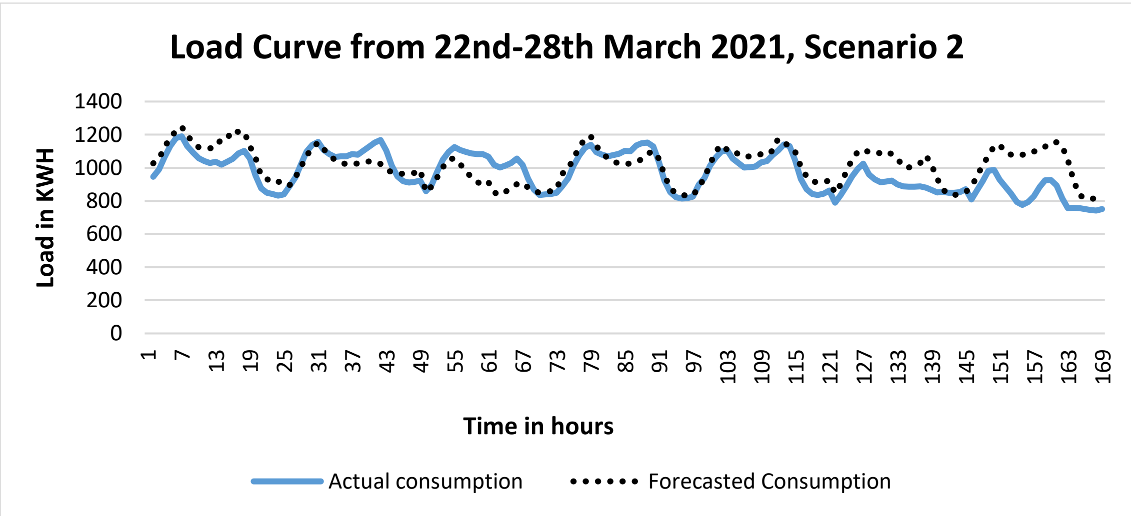

Figure 14. Actual and forecasted load consumption for scenario 2

In some conditions for forecasting, previous day data is not available in utility companies, so keeping that situation as the frame of reference, in this research, two days' previous data is used for electrical load forecasting, that was available on the NordPool website, finally after forecasting the consumption and MAPE is calculated by following formula, which is given below,

$$MAPE = \frac{1}{n} * \sum_{i=1}^{n} . \frac{Actual\ value - Predicted\ Value}{Actual\ Value}$$

Where,

n is the total number of time stamps, we are including in our forecasting

Table 9. MAPE calculation from 22 march to 28 march using two days' previous load data.

Days of week	**MAPE**
Monday	9.6 %
Tuesday	4.35 %
Wednesday	8.4 %
Thursday	4 %
Friday	4.7 %
Saturday	11 %
Sunday	20 %

In this experiment, the author successfully decreased the forecasting error for Monday from 11% to 9.6% that can be seen in figure 14. However, another problem that arises in this experiment was that the forecasting error jumped to 20% for Sunday, which is not acceptable in any scenario; it is very high.

To overcome this problem, the author utilized a different approach; the author used last week's consumption data as target for ANN and computed the error again in the third experiment.

4.3 Scenario III

In the third Scenario, forecasting is being done by using one-week, previous data. The forecasting is improved a little, with the lowest forecasting error of 2.9% for 23rd March, as shown in table 13.

The initial input parameters of ANN model used in experiment 3 are shown in table 10 and training parameter of the network are shown in table 11.

Table 10. The primary inputs of ANN, using the target inputs from previous load consumption of one week.

Inputs of neural network	Number of hours of a day Temperature of the day Pressure of the day Month of the year Humidity Day of the week
Number of neurons in hidden layer	10
Activation function	tanh
Number of neurons in output layer	1
Activation function of output layer	Purelin (Linear)

Table 11. shows the training parameters for third experiment

Number of Epochs	300
Training data	15%
Validation of data	15%
Error estimator	MAPE
Testing Input	1 January-18 March 2021

Table12, shows the hourly based forecast results for complete cycle of 24 hours of the day, from Monday 22nd march to Sunday 28th march. After getting the forecasted results mean absolute percentage error was computed for each day starting from Monday to Sunday and MAPE results for experiment 3 are shown in table 13.

Table 12. forecasting from 22nd to 28th march using one-week previous load consumption data sets.

Time	**Monday**		**Tuesday**		**Wednesday**		**Thursday**		**Friday**		**Saturday**		**Sunday**	
Hrs.	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**
23	946	926.166	889	914.2397	859	902.9738	889	947.4507	897	1017.275	790	890.0566	809	853.3973
22	991	974.09	942	972.4478	898	969.6348	936	999.7232	938	1055.001	834	935.2479	863	886.083
21	1063	1050.169	1020	1036.4579	977	1042.941	1014	1077.583	1011	1123.283	887	985.4191	919	940.0881
20	1131	1136.929	1098	1101.8349	1050	1108.9311	1076	1137.3031	1062	1087.961	947	1049.988	983	1000.242
19	1180	1211.382	1138	1154.1376	1097	1164.4498	1121	1188.3124	1098	1201.301	995	1087.9403	987	1031.562
18	1190	1200.184	1157	1189.8978	1126	1184.9588	1138	1207.3579	1106	1258.03	1024	1097.5074	928	1071.783
17	1132	1128.062	1107	1148.2449	1107	1161.9055	1095	1188.329	1055	1204.768	960	1080.9736	886	1234.244
16	1093	1135.43	1082	1103.4422	1097	1132.6339	1079	1175.1926	1028	1171.788	930	1055.9333	843	1163.189
15	1057	1122.207	1065	1090.5112	1086	1120.9816	1070	1183.9859	1001	1147.92	913	1045.8842	794	935.6182
14	1040	1137.256	1069	1078.3084	1083	1139.9773	1078	1189.0457	1003	1128.717	916	1033.5686	775	913.3716
13	1028	1131.424	1070	1095.1196	1082	1151.3478	1084	1183.8777	1007	1133.277	922	1017.9177	794	891.2992
12	1037	1130.411	1083	1111.6558	1067	1170.5057	1102	1158.579	1032	1132.252	900	1023.2298	829	858.4814
11	1020	1124.422	1079	1120.593	1018	1191.7965	1101	1145.8183	1040	1148.93	888	1031.2625	884	860.5819
10	1036	1158.667	1104	1158.5157	1001	1194.8072	1135	1151.6908	1078	1181.434	885	1028.8971	924	842.2799
9	1053	1235.31	1127	1188.8406	1014	1225.0007	1148	1160.6399	1105	1192.09	886	1029.5543	926	853.9755
8	1086	1230.101	1153	1207.8046	1029	1228.6604	1153	1184.8776	1140	1225.906	887	1028.6982	893	901.1569
7	1102	1195.919	1168	1183.9661	1055	1187.5556	1130	1181.635	1132	1233.176	880	1021.1604	815	902.0243

6	1055	1083.986	1105	1087.889	1018	1091.9782	1033	1120.0466	1047	1142.438	866	1015.9592	757	871.1891
5	955	932.1953	1014	981.2688	935	978.6311	923	1010.9185	932	1038.883	852	983.2523	758	852.2119
4	877	863.4364	948	905.9993	869	904.5961	854	918.3618	871	959.8067	856	960.7545	757	850.085
3	849	819.8246	919	879.4411	835	874.0922	822	868.6928	842	923.7072	850	957.6276	750	851.1495
2	842	814.5699	911	875.8205	839	863.8713	815	857.5999	836	912.963	849	960.202	745	854.2279
1	832	818.7566	914	877.8387	842	864.8628	817	866.5739	844	921.4597	854	968.0947	742	859.7239
0	840	860.8156	923	892.553	850	868.8255	826	884.1387	863	957.7509	871	984.8896	750	864.3735

The graphical representation of table 12 shown in figure 15.

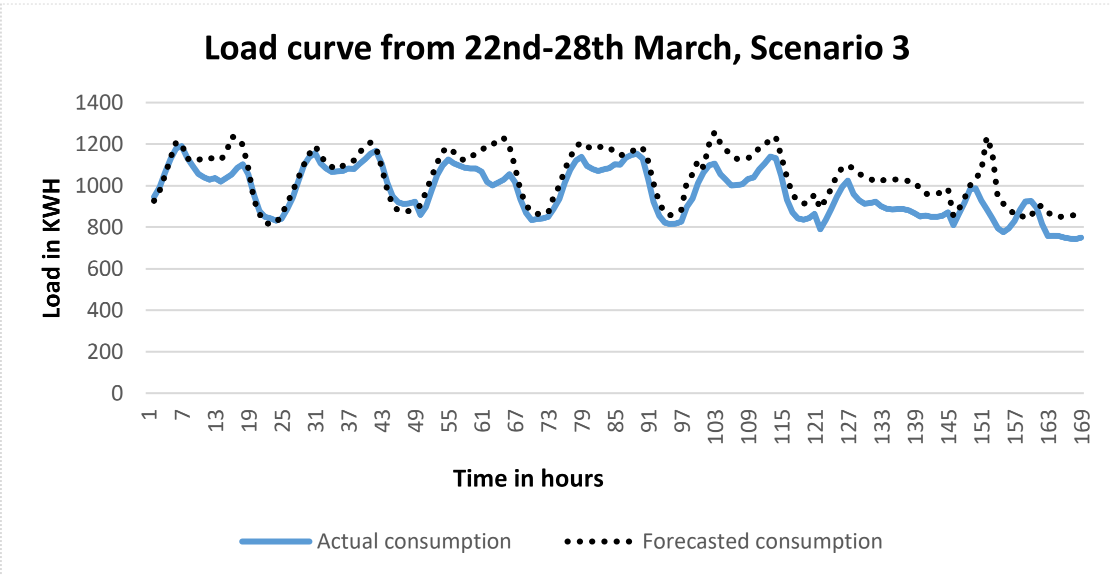

Figure 15. Actual and forecasted load comparison in scenario 3

It can be seen in figure 15, that from the start of the week on 22nd March the actual and forecasted values are very close to each other and MAPE is very small in scenario 3, as compared to other first two scenarios. Forecasting error remained around 10%, average on weekends. Overall performance of scenario 3 is better than first two but still not low, according to our expectation.

The following formula is used to calculate the forecasting error.

$$MAPE = \frac{1}{n} * \sum_{i=1}^{n} . \frac{Actual\ value - Predicted\ Value}{Actual\ Value}$$

Where,

n is the total number of time stamps, we are including in our forecasting

After calculating the mean absolute error for all days of the week by using the formula shown above, results are shown in table 13.

Table 13. The results of mean absolute percentage error.

Days of week	**MAPE**
Monday	5.2 %
Tuesday	2.97 %
Wednesday	7.8 %
Thursday	6.4 %
Friday	10.2 %
Saturday	13.2 %
Sunday	11 %

The good point in the third experiment was that the forecasting error on Tuesday is 2.97% which is the lowest achieved error in all three experiments.

The failing reason for this experiment was again the error 11% MAPE on Sunday, 13.2% MAPE on Saturday. That is very high, and does not make any sense of the accuracy of ANN.

Therefore, it was decided to increase the dataset values, and those have proceeded in experiment 4.

4.4 Scenario IV

It is clear from the first three experiments in scenario I, scenario II and scenario III that, mean absolute percentage error is high, initially at the start of the week (Monday) and then remains stable to some extent but unfortunately increases 13%,11%, and 20% on weekends. Therefore, different approaches with different algorithms were utilized in this research but did not see any good improvement in the error accuracy of forecasting.

To overcome this problem, it is suggested to increase the training dataset values so that ANN has enough data to learn; previously, data sets being used in experiments were two and half months. (1^{st} January – 18^{th} March 2021). After that, the range of dataset values increased to one year three months from 1^{st} January 2020 to 18^{th} march 2021, and all experiments of one-week short-term load forecasting performed again; there was a significant boost noticed in forecasting accuracy that is shown in table 17.

Initial parameters of the selected model are shown in table 14,

Table 14. Inputs of the network with one year three months' data sets values

Inputs of neural network	Number of hours of a day Temperature of the day Pressure of the day Month of the year Humidity Day of the week
Number of neurons in hidden layer	20
Activation function	tanh
Number of neurons in output layer	1
Activation function of output layer	Purelin (Linear)

After finalizing the inputs and creating a model, training parameters are selected which are shown in table 15,

Table 15. shows the training parameters used in experimental model of ANN

Parameters of model	**Values**
Number of Epochs	300
Training data	15%
Validation of data	15%
Error estimator	MAPE
Testing Input	1 January 2020 to18 March 2021

Table16, shows the hourly based forecast results for complete cycle of 24 hours of the day, from Monday 22nd march to Sunday 28th march 2021. Using one year and three months' dataset values starting from 1st January 2020 to 18th march 2021. After getting, the forecasted results mean absolute percentage error was computed for each day starting from Monday to Sunday and MAPE results for experiment 4 are shown in table 17.

Table 16.Forecast from 22nd march Monday to 28th march Sunday with one year three months data set values

Time	**Monday**		**Tuesday**		**Wednesday**		**Thursday**		**Friday**		**Saturday**		**Sunday**	
Hrs.	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**	**A Kwh**	**F Kwh**
23	946	937.2235	889	948.6691	859	844.8761	889	882.7844	897	888.2745	790	803.8644	809	762.0814
22	991	987.5714	942	977.1159	898	942.644	936	979.6624	938	933.0012	834	901.0062	863	833.7127
21	1063	1063.291	1020	1033.852	977	1020.876	1014	1046.42	1011	1022.227	887	987.3354	919	898.967
20	1131	1133.496	1098	1132.076	1050	1098.179	1076	1089.04	1062	1074.06	947	1007.516	983	939.8656
19	1180	1193.586	1138	1183.326	1097	1138.988	1121	1128.691	1098	1115.046	995	1012.431	987	953.4023
18	1190	1196.038	1157	1183.167	1126	1137.556	1138	1147.225	1106	1131.678	1024	1085.199	928	970.8008

17	1132	1151.984	1107	1134.812	1107	1132.859	1095	1119.594	1055	1111.042	960	999.9999	886	916.2358
16	1093	1115.954	1082	1088.702	1097	1082.551	1079	1072.011	1028	1075.044	930	911.8761	843	833.376
15	1057	1122.406	1065	1059.661	1086	1004.965	1070	1055.263	1001	1075.477	913	952.611	794	881.216
14	1040	1144.248	1069	1069.161	1083	980.8971	1078	1064.286	1003	1077.159	916	968.8746	775	830.3268
13	1028	1124.183	1070	1026.316	1082	1022.228	1084	1060.331	1007	1083.582	922	1007.877	794	818.3711
12	1037	1126.337	1083	1101.078	1067	1083.891	1102	1074.794	1032	1107.609	900	1012.56	829	840.717
11	1020	1155.738	1079	1080.57	1018	1029.653	1101	1070.902	1040	1194.233	888	1033.012	884	871.5046
10	1036	1151.174	1104	1087.049	1001	1085.641	1135	1030.471	1078	1173.319	885	1017.897	924	860.8795
9	1053	1163.437	1127	1147.565	1014	1062.755	1148	1095.368	1105	1122.425	886	1016.731	926	841.2541
8	1086	1108.436	1153	1133.455	1029	1182.759	1153	1144.837	1140	1157.038	887	1113.877	893	842.2117
7	1102	1080.922	1168	1198.059	1055	1168.877	1130	1159.288	1132	1132.499	880	1159.425	815	900.2289
6	1055	1041.551	1105	1084.667	1018	1083.457	1033	1097.242	1047	1066.876	866	947.498	757	923.1076
5	955	941.6882	1014	919.6968	935	914.765	923	1008.128	932	924.4214	852	932.885	758	908.1046
4	877	908.4611	948	889.9678	869	965.934	854	939.4823	871	858.4734	856	844.677	757	748.2322
3	849	887.0193	919	864.3375	835	935.907	822	916.3324	842	822.4636	850	848.961	750	778.4274
2	842	817.5871	911	778.398	839	911.541	815	866.2295	836	815.7247	849	811.6896	745	735.1814
1	832	843.7948	914	859.5195	842	814.781	817	842.2987	844	817.6003	854	852.719	742	786.4587
0	840	846.8429	923	843.4169	850	819.5133	826	821.541	863	824.9818	871	872.3416	750	779.7214

The graphical representation of table 16 are shown in figure 16.

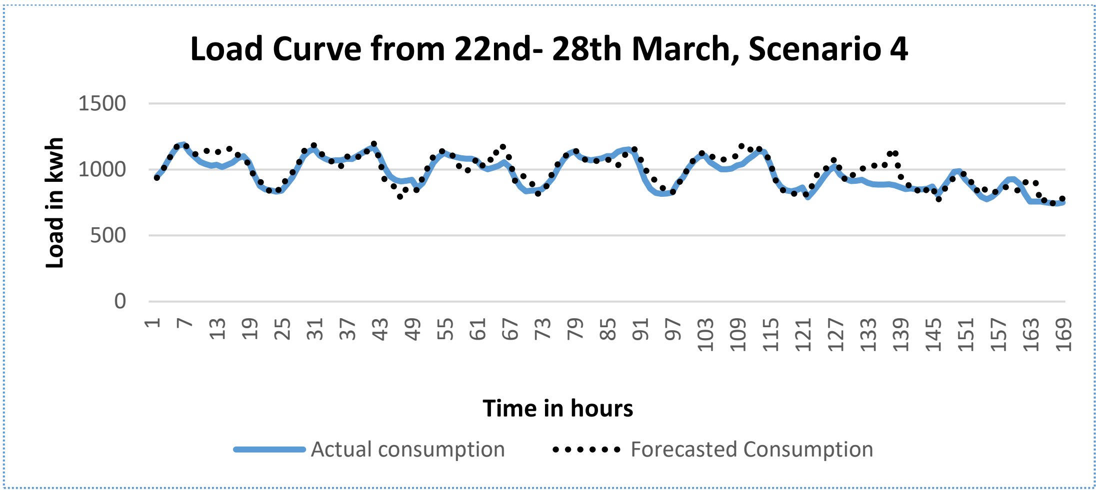

Figure 16. Actual and forecasted consumption for scenario 4

It is noticed that in scenario 4, results are the best achieved result of this research work, starting from 22nd March, the forecasted values are following exactly the same pattern as actual electrical consumption, a difference can be seen in figure 16. Even on weekends, MAPE for Sunday is 5.53%, which is the lowest achieved forecasted error on weekend.

After obtaining all the forecasting results of one week, forecasting error was calculated using mean absolute percentage error, that is shown in table 17.

Table 17. MAPE error of forecasting using data set of 1 year 3 months' data set values

Days of week	MAPE
Monday	3.99%
Tuesday	3.77%
Wednesday	5.4 %
Thursday	3.67 %
Friday	3.77 %
Saturday	8 %
Sunday	5.53 %

One thing is clear from the results of scenario 4 that, the more data is provided to artificial neural networks for training, testing and validation, accordingly much precisely they forecast. That can be seen in the MAPE results shown in table 17. Some significant improvements are, previously MAPE for Monday was 11% but in scenario 4, it is reduced to 3.99%. MAPE on Saturday in first three scenarios was 12% but in scenario 4 it is reduced to 8%. MAPE on Sunday in first three scenarios was 8.7% but in scenario 4, it is reduced to 5.53%. Therefore, scenario 4 is the most efficient in all of the experiments done in this research work.

4.5 Forecasting Error on Holidays

In forecasting electricity consumption on holidays, forecasting error seems very high because the electricity consumption on holidays is not the same as normal working days. One reason is that mostly the commercial activities in schools, colleges, universities, and tech companies are closed on weekends, and Neural Network learns the unusual behavior and forecasts electricity.

One solution to this problem, the electrical consumption data should be collected from the utility company about the type of their customers like (residential, agriculture, business or commercial) and baseline is set according to all commercial activities that are not working on holidays, and this non-operating load can be subtracted from the forecasted values by ANN. e.g., In experiments of this research for NordPool, it is observed that at the start of the weekend (Saturday) always around 1000kwh load is cutoff from total demand keeping 500kwh load tolerance; therefore, it is suggested that, on Saturday, non-operational load that is around 1000kwh should be subtracted from the forecasting values of ANN. This will greatly increase the accuracy of the load forecast.

For 27th march 2021, using the previous day data approach, the forecasted and actual values were,

$$Actual\ Value = 21442kwh$$

$$Forecasted\ value = 23102.86kwh$$

$$MAPE = 8\%$$

But after subtracting the fixed load, that is around 1000kwh, is not operational during holidays. This fixed load is observed from historical dataset values.

$Actual\ Value = 21442kwh$

$P' = Forecasted\ value - fix\ load(not\ in\ operation\ on\ holidays)$

$P' = 23102.86kwh - 1000kwh$

$P' = 22102kwh$

Where,

P' represent the modified forecasted value

So, by using the above values we can reduce our forecasting error, the MAPE on 27th march 2021 is 5%.

4.6 Effect of temperature on forecasting

In order to check that temperature is playing an important role in forecasting, temperature data is being removed from the inputs and train the network with other inputs like hours, day of week, pressure, number of months and humidity. There is a significant increase in forecasting error, which can be seen in table18.

The worst-case occurred when the model error increased from 5.4% to 6.7%. So, it is very obvious that temperature is playing a very important role in forecasting.

The removal of temperature from the input vector significantly degrades the quality of forecasting; for this experiment, the random data sets from 22nd March 2021 to 28th March 2021 are used.

Table 18. Effect of temperature on forecasting

Epochs	**Forecasted error with temperature**	**Forecasted error without temperature**
25	3.99%	4.34%
25	3.77%	4.2%
25	5.4%	6.70%
25	3.67%	3.94 %

As in these experiments, the forecasted temperature dataset values are used; one problem in forecasted temperature is that it's not always 100% accurate. So, if the temperature varies, it will affect the forecasting accuracy. In this research, many experiments are performed on 23rd March and 24th March 2021, to show how forecasting error will increase with shift in temperature, the results are shown in table 19 and table 20.

Table 19. showing how forecasted error varies with temperature shift (±2 degrees)

Day	**Forecasted error**	**Forecasted error with temperature shift +2 degrees**	**Forecasted error with temperature shift -2 degrees**
23rd march 2021	3.77%	4.74%	4.61%
24th march 2021	5.4%	6.41%	6.20%

In table 19, it can be seen, that if the temperature shifts, it causes around 0.9% increase in electrical forecasted error, like if the forecasted error on 23rd March was 3.77%. It increased to 4.74%, when temperature shift occurs +2 degrees and when the temperature shifts -2 degrees, it still causes the increase in forecasted error, which is 4.61%. In both cases (±2 degrees shift), the forecasted error is comparatively less in -2 degrees shift.

Table 20, shows the temperature shift of ±1 degrees Celsius effect on forecasting accuracy.

Table 20. Forecasted error varies with shift in temperature (±1 degrees)

Day	**Forecasted error**	**Forecasted error with temperature shift +1 degrees**	**Forecasted error with temperature shift -1 degrees**
23rd march 2021	3.77%	4.97%	4.47%
24th march 2021	5.4%	5.91%	5.73%

From table 19 and table 20, it can be seen that if the forecasted temperature varies just +1 degree or -1 degrees Celsius, then the electrical consumption forecasted error also increases more or less 0.7% on average, So one percent error in electrical consumption forecasting results in the loss of millions of euros.

In order to forecast the consumption of electricity more precisely, more accurate forecasted temperature data is needed. If the forecasted temperature is almost the same as actual day temperature, then the electrical consumption forecast error will also be very low; from these experiments, it is concluded that forecasting of electrical consumption depends upon several factors where forecasted temperature plays an important role

4.7 Data Rescaling

In this research, five inputs are used: temperature, pressure, humidity, day of the week, and 24-hour system data. Inputs are rescaled to see their impact on forecasting accuracy; Therefore, the temperature values are changed by adding temperature shifts tolerances. It is observed that any temperature change will greatly affect the forecasting accuracy. Further in this research, other values are rescaled, like values of pressure, which do not show any significant impact on forecasting accuracy. Pressure values were not changed that much in the forecasting period, so there is no need to rescale them; they remained overall between 750-765mm Hg. In the last, the number of hours per day was rescaled into their sine representation and cosine representation.

4.8 Network Accuracy through K-fold cross-validation

K-fold cross-validation is an evaluation process in which the accuracy of the model is being tested, like how well the model is predicting, is it predicting good for specific data sets values, or what will be the behavior of the model for other data set values (unknown data set or new data set values).

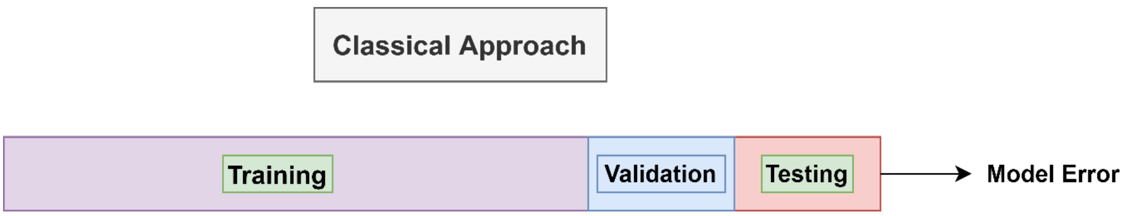

Figure 16. Classical approach for calculating model error in ANN

In the classical approach, the data set is divided into three steps, training, validation, and testing as shown in figure 16. Therefore, dividing the data into three sets the data set is reduced, that is being used for training the model. Accordingly, if the network is trained with a specific amount of data, the model may behave very well in forecasting targets for these inputs data set values. Conversely, maybe if a new set of data inputs are applied to the network, it shows very bad performance in forecasting targets. Here is the solution for this problem to check how the model will behave for unknown dataset values that are never subject to it before; that solution is K-fold cross-validation.

In k-fold cross-validation, the data is being split into k sets, which are used for training the model. K-fold cross-validation allows using of all the available data for training purposes of the network. So, it's no longer necessary to keep the testing data completely apart.

By choosing a different set of training data, model performance seems to increase because the model is trained on all data set values. Instead of depending upon a specific training set. Cross-validation gives a more stable estimate of how the system behaves overall for different dataset values. The figure 17, shows the k-fold cross-validation approach to estimate the accuracy of the model.

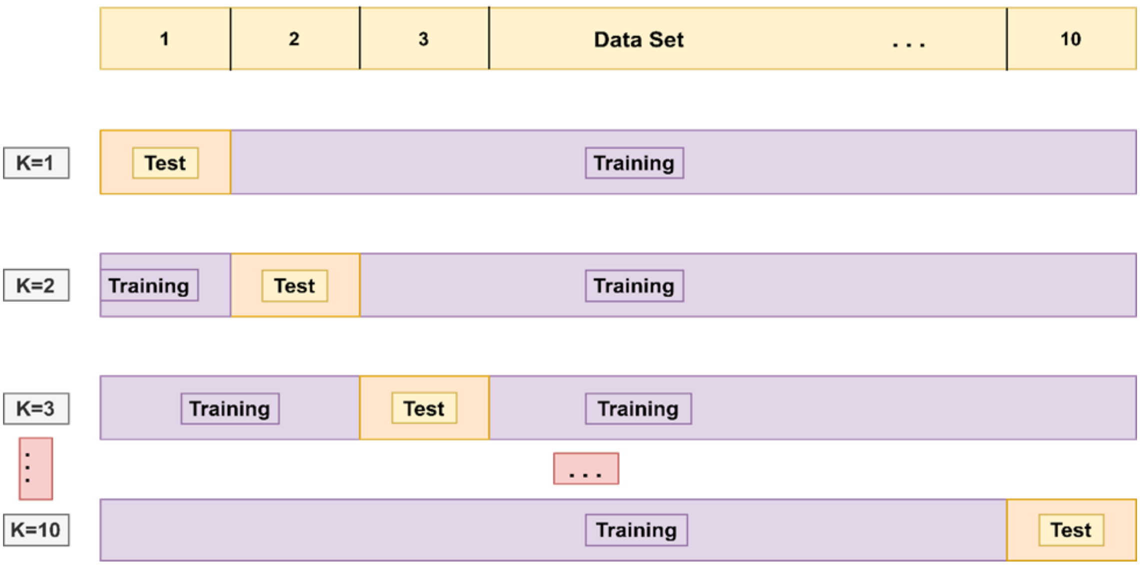

Figure 17. K-fold cross validation to check the model accuracy using 10 folds

In this experiment, 10-fold have been chosen for cross-validation. Therefore, all the test data is being divided into 10-sets. So first of all, 80% of the data is used for training and 20 percent of data as testing. These experiments are repeated ten times, each time selecting a different data set. After performing different iterations on data set values and switching the training fold so that all folds' data should include at least once in training.

When this cycle completes, there are ten accuracies values, one for each model as a result of k-fold cross-validation, finally, by calculating the average of all accuracies and estimate the overall performance of the model used in this research.

$$Model\ Accuracy = Average\ Accuracy(\sum_{i=1}^{n}(All\ k\ fold\ models\ accuracies))$$

$$Model\ error = \frac{1}{10} * (\ -34.28 - 13.12 - 24.61 + 27.15 + 25.55 - 10.12 + 24.50 + 44.10 - 4.79 \quad - 78.990)$$

$$Model\ error = 4.461\%$$

The selection of several folds must be very precise so that the testing fold has enough data to accurately estimate the model's efficiency. Therefore, the testing fold would never be too small.

Cross-validation is used to choose the model hyper-parameter to train the network more efficiently. Therefore, to be sure that this designed model will behave accurately for unseen or new data presented to it.

Table 21. Different set of data inputs provided to network to investigate the model error

First Inputs (6-Inputs)	**Second Inputs (11-Inputs)**
Month of the year Day of week (numeric number) The average temperature of Tallinn Humidity of Tallinn Average Pressure of Tallinn Hours of the day (24 hours)	Month of the year Cosine reflection of month number Sine reflection of month number Cosine reflection of 24-hour system Sine reflection of 24-hour system Day of the week Sine reflection of day of week Cosine reflection of day of week

	Average Pressure of Tallinn Humidity of Tallinn Average temperature of Tallinn

In this part of the research, experiments are performed using the new inputs that are 13 in total and compare the results with the previous results obtained from 6 inputs. The results are shown in Table 22. It is very clear that if we increase the number of inputs, better ANN will learn and generates output with less forecasting error.

Table 22, showing the model error accuracies for different sets of data inputs to ANN on 23rd March.

Table 22. showing the model error accuracies for different set of data inputs to ANN on 23 march.

Error in forecasting	**Error with 6 inputs**	**Error with 11 inputs**
Model error	4.46%	4.39%

Prediction of electrical consumption is not an easy task, and we cannot predict exactly what is the actual behaviour of electricity consumption. Therefore, it is a very difficult task to forecast the electricity consumption with the highest precision.

In this research, the author has considered the latest research paper published in 2020 [42] and compared the forecasting results, the forecasted results in this thesis shows very significant improvement. Overall experiments on STLF in this thesis were more feasible with less mean absolute percentage error than the referenced researched article.

As we all know, ANN sometimes never gives good accuracy in forecasting when we train them at first, so we need to train them several times until we get our desire outcome. Therefore, a program was made so that anyone can obtain the results from the trained ANN network and store results in a database; if any un-professional person wants to see the forecast of any day, he can press the number of that day and see all results without training the ANN.

5 Numerical Results

In this chapter, the results are concluded based on the analysis of different scenarios performed in the chapter 4, after the successful experiments of simulations in MATLAB through Artificial Neural Network to forecast short-term load consumption in Estonia.

Implementation of different approaches in first three scenarios to minimizing the forecasting error for STLF and by using historical test data set from 1st January to 18th march 2021. Initially experiments are performed by increasing and decreasing the number of epochs, number of neurons in hidden layers, with different transfer functions until satisfactory outcomes obtained, then finally the best model is selected that was best suited for load forecasting with minimum forecasting error. The resulted forecasted output of first three scenarios are shown in figure 18.

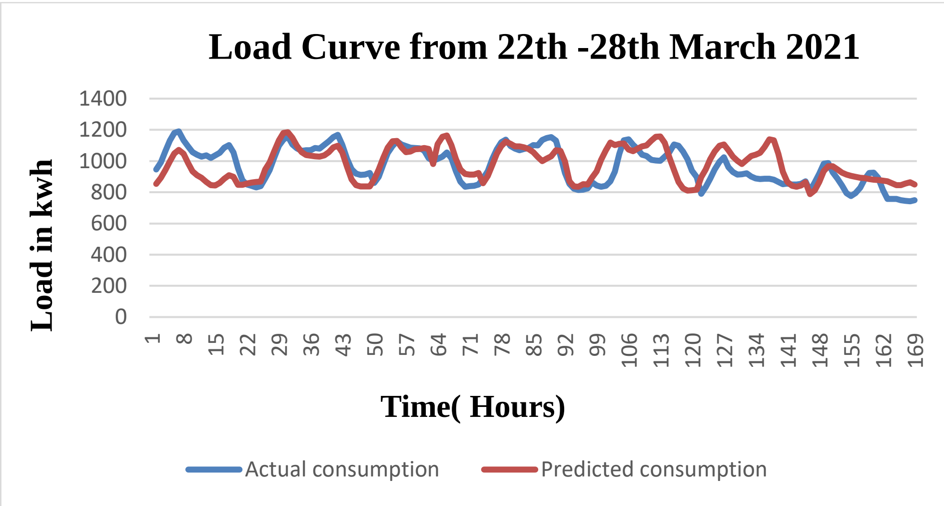

Figure 18. Shows the actual electrical consumption and forecasted electrical consumption for one week in Estonia

In these experiments, MS Excel 2016 is used to plot the relationship between daily actual electrical consumption in Estonia with the forecasted electrical consumption from 22nd March to 28th March, which is obtained from our ANN model.

The redline represents the predicted value, while the blue line represents the actual values of electrical load consumption. From the graph, it can be seen that predicted values are a bit smaller than the actual values on Monday and forecasted error is high that is around 11%. After that, from Tuesday (hour 23 onward), our forecasted values are very close to actual values, even at some points both forecasted values and actual values overlap with each other; actually, those are the best points of forecasting, where our network forecasted very accurately (actual values).

Then from Friday (125th hour), the error again starts to rise and is followed by Saturday and Sunday. The major difference occurs after Friday when the weekends start. During Saturday and Sunday, the commercial load is switched off, and the residential load contributes maximum electricity consumption. Therefore, from Saturday, the difference in load drop calculated by subtracting the forecasted load value from the actual load value lies between 1000kwh to 1500kwh.

The same pattern will be followed again starting from next week, on Monday almost 1000kwh load is added to the electricity consumption system.

The first three scenarios fail because of high forecasting error on weekends, so it was decided to increase the historical data set values. After that significant improvement in forecasting accuracy observed in scenario 4.

In this research work results are verified with a published research article on short-term load forecasting by [42] based on Artificial Neural Network. In the first stage, the forecasted error was calculated from the referenced research article and validated with this research work. In the second stage, a MATLAB simulation setup was created based on Artificial Neural Network to forecast the short-term electrical load consumption, which resulted in a high forecasting error. The error was further reduced in **scenario 4**, by iterative improvements, e.g., introducing multiple inputs and optimization of the ANN model, which resulted in the forecast error on weekends to be as low as 5.5%, whereas on weekdays as 3.77%. The input data set used for scenario 4 was from 1st January 2020 to 18th March 2021.

The best forecasted results of scenario 4, with minimum mean absolute percentage error are shown in figure below.

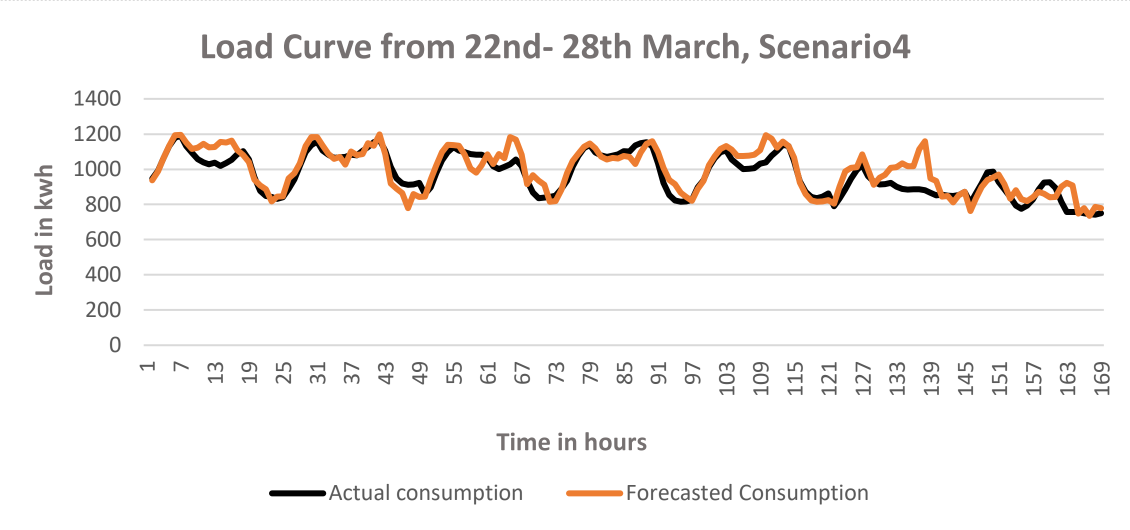

Figure 19. One-week complete actual load plotted against forecasted load consumption

In figure 19, it can be seen from the start of week on 22nd March the forecasted values lie exactly on actual consumption values. Both curves follow the same pattern of electrical consumption with a minute difference on weekends.

The reduction in mean absolute percentage error becomes possible when the number of inputs to the model's increases. Therefore, it was realized that increasing the number of inputs results in better learning of ANN. When the ANN has enough data set of learning, testing, and validation, it performs much better forecasting. Secondly, the author would like to mention that the inputs should not be chosen randomly. Instead, we should be very precise in selecting inputs by considering which data sets will be effective for forecasting.

5.1 Forecasting Evaluation

For evaluation purposes, a graph was plotted between the actual consumption and prediction by ANN. Mostly it was observed that, high forecasting error occur on Monday when the week starts and at the end of the week, Sunday. Therefore, a graph was plotted for Monday, in figure 20.

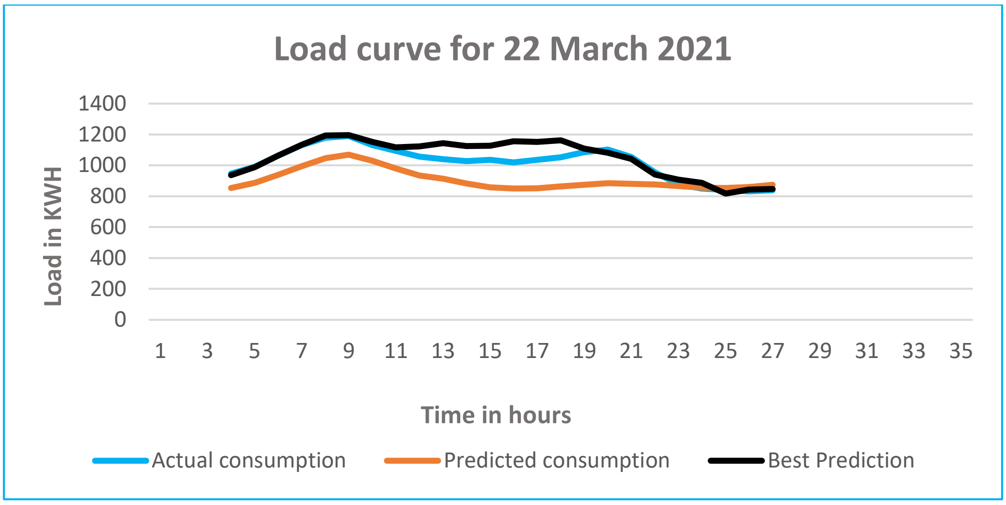

Figure 20. comparison of results from first three scenarios and 4th scenario

Blue curve represents the actual consumption, **orange curve** represents the predicted values (forecasted values) obtained from first three scenarios and **black curve** represent the predicted values obtained from scenario 4.

It can be seen, that the black curve is showing the best prediction in comparison with the actual values and with very small mean absolute percentage error of 3.99%. So it is concluded that by increasing the input data set values the accuracy of ANN increases.

Second graph was plotted for Sunday 28th March. As shown in figure 21.

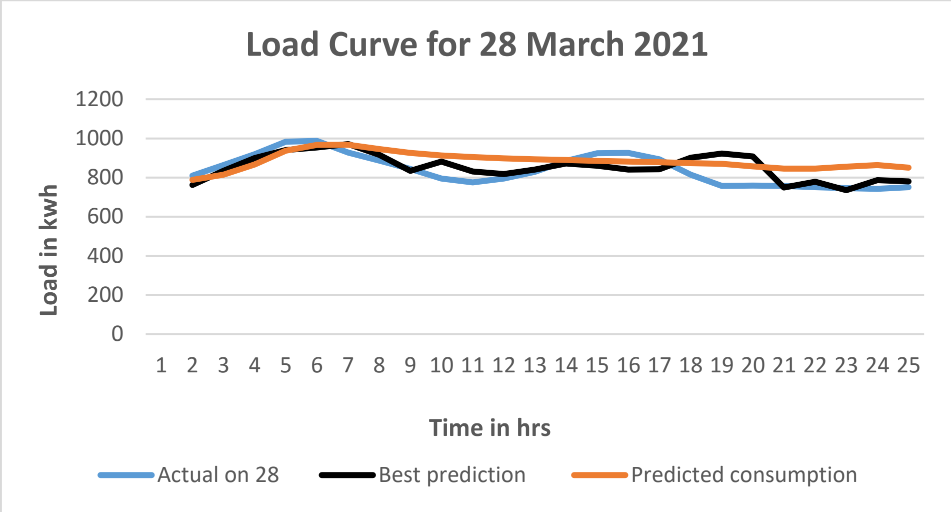

Figure 21. comparison of results from first three scenarios and 4th scenario

The black curve which is obtain from scenario 4, again showing much improved performance, from the start to end with small mean absolute percentage error of 5.5% as compared to orange line which MAPE is 8%. Therefore, we are successful in decreasing the MAPE on Sunday.

While on the rest of working days the mean absolute percentage error is very small, the main difference occurred on weekends and Monday which are shown above in figure 20 and figure 21, with improved accuracy.

One solution to this increased forecasting error is that we can subtract the fixed load that is being shut down during weekends.

5.2 Comparison of Forecast between all four scenarios

Depending upon different conditions and circumstances of utility companies and consumers, experiments have been performed by considering one day before actual consumption, two days before actual consumption, and finally one week before actual consumption of electricity and concluded the results in table 23.

Table 23. Final conclusive comparison of MAPE for whole week in all four experiments

Days of week	Scenario I	Scenario II	Scenario III	Scenario IV
Monday	7.9%	**8.66%**	5.2%	**3.99%**
Tuesday	4.30%	**4.35%**	**2.97%**	3.77%
Wednesday	5.4%	**8.40%**	7.8%	**5.4%**
Thursday	3.04%	4%	**6.4%**	**3.67%**
Friday	4.2%	4.7%	**10.20%**	**3.77%**
Saturday	12%	11.39%	**13.2%**	**8%**
Sunday	8.7%	20%	**12.12%**	**5.53%**

Comparing the results of all four forecasts, the highest and lowest MAPE are highlighted with Bold font, highest MAPE value represent the worst forecast and lowest MAPE represent the best forecast. Therefore, it is clear that the best results lie in scenario 4. It is concluded that the forecasting error was different for different approaches. The forecasting error varies from 5.2% to 7.9% then 8.88% for the test data set from 1st January 2020 to 18th March 2020. There was a significant improvement in accuracy that is 3.99% MAPE on Monday, observed when the historical data set values increased to one year and three months from 1st January 2020 to 18th march 2021. Secondly, during the working days from Tuesday, Thursday, and Friday the forecasting error is not that much high but when the weekend starts again, degradation in forecasting performance increases.

Before MAPE on Monday was 8.66% which is decreased to 3.99%. On Tuesday MAPE was 4.55% which was reduced to 3.77%, MAPE reduced in a same pattern in the following working days Wednesday and Friday. The main improvement observed on Saturday when MAPE was reduced from 13.2% to 8%.

5.3 Effect of weather on forecasting

In the area under observation for short-term load forecasting, weather plays an important role. There are many seasons during a year, mainly winters, summers and a very short spring in Estonia. During the winter, temperature in Estonia drops below zero as an average, and sometimes it is very severe and drops under minus 25 degrees Celsius. Therefore, electrical consumption in those days is very high because of the heating requirement in all the residential and commercial buildings.

In summers when the temperature is around 20 degrees, mostly consumption has been resulted from Air-conditions. There is a mid-season called spring when the weather is not that cold and not that much warm; in other words, we can say pleasant weather; at this period, electrical consumption is very low compared to rest of year consumption.

All these transitions of season affect consumption of electricity, which in return affect the forecast, that is our main topic of discussion. It has been shown in experiments (section 4.5) that how a small change in temperature affects the forecasting.

Other factors play an important role quite similar to temperature: atmospheric pressure and humidity. When the temperature is high, the air contains more water molecules, and humidity is very low at that moment; likewise, when the humidity is high, the temperature of the air reduces.

The atmospheric pressure is high during summers when there are no clouds and full sunshine, and it is low during cloudy days or rainy days. Therefore, atmospheric pressure is also an essential factor in consideration for electricity forecasting.

5.4 Daylight influence

Another important factor that affects electric consumption is the daylight which is also essential because during winters in Estonia, there is no daylight, mostly its dark so there are fewer folks outside, even in their free time, they prefer to stay at home, which causes a major increase in electricity consumption.

On the other hand, in summers, the outside activities of folks increase, and they spend more time out of their houses; therefore, electricity consumption decreases. Finally, we can say that daylight has a direct relation with electricity consumption.

5.5 Model inputs

At the beginning of these forecasting experiments, when a random input in the form of the number of months, days of week, and hour of the day were selected, the forecasting error was around 20%. Which was later practiced to train the network with three months' data sets, but the accuracy was not improved. In the second attempt, the parameters that directly affected the consumption of electricity, like temperature, atmospheric pressure, and humidity, were included. After that, these new inputs were subjected to the network that provided a very significant improvement in the accuracy of this network.

One of the major issues in training this model was applying the previous day's temperature to forecast the electrical consumption for the next day. When the temperature data for the same day provided by weather agencies were utilized for the electricity forecasting, the accuracy improved.

In the final experiment, when the input data was subdivided, like sine reflection of month number and cosine reflection of month number, the new input data sets increased to 11, which were previously 6. These increased inputs have also contributed to increasing forecasting accuracy to 0.07%, shown in table 21, section 4.7.

Hence, we can conclude that the number of inputs is proportional to the forecasting accuracy. With an increase in the number of inputs, the ANN trains better and provides good solutions.

5.6 Problems associated with input data sets

The problem faced during the collection of data set was that the weather data obtained from the rp5 website, very often there was missing data of temperature, pressure, and humidity for some hours, Therefore, it is an uncertain situation to predict for those hours by the way that interpolation methods could formulate missing temperature data sets when single values were missing. However, when four or five values were missing in a column in historical data, the average has been taken from previous values substituted in blank spaces.

5.7 Further Development

It is seen that the suggested model's performance degradation, in this dissertation increases on holidays and sometimes during the working days. There are some other factors that are affecting the performance of ANN rather than inputs. By finding the root cause of deviations in forecasting, this model can be improved more, e.g., by introducing some different pre-processing processes for further improvement of ANN accuracy.

The model ensemble is a process in which different models are combined to achieve more precision in a network. Therefore, an ensemble model should be utilized to perform these computations for electricity forecasting. However, due to computational resources and time limits, this method could not be applied in this case.

A better approach in the future will be selecting different models and use them for experimentation and finally take an average output, the same as a hybrid approach. It is believed that the hybrid models will be considered in the future for the computations of electricity forecasting.

5.8 Conclusion

- ANN shows very different behavior as compared to linear regression models. If there is a slight change in input parameters, it will degrade their performance.

- ANN needs several data inputs for training, and the more data is used for training the ANN, the better performance can be expected in forecasting. One of the major issues is that due to global warming, data provided to train the ANN is limited, e.g., 2020 in Estonia, the temperature in March was around 15 degrees, but this year it could not rise to 8 degrees. According to ERR, 2020 was the warmest year in the history of Estonia [13].

- Not just the input data sets affect the ANN performance but, the way they are being subjected to ANN is essential, e.g., when the number of inputs increases by adding their sine and coefficients, accuracy improves to some extent.

- If the weather temperature forecast deviates from its real values, it also affects the forecasting accuracy of ANN.

- Optimally choosing the hyperparameter of ANN can only increase the performance.

- In different studies, it is clear that the new trend of forecasting is moving towards hybrid approaches because a single model can never have more accuracy than hybrid models.

- To make ANN more user-friendly, a program is made in MATLAB and it has ability to store in it the best-trained model results like predicted values, forecasting error, and plot of actual and predicted values.

- A generalization test was made using k-fold cross-validation to show that the model used in this research study generates very stable results.

References

[1] “Forms of energy - U.S. Energy Information Administration (EIA).” https://www.eia.gov/energyexplained/what-is-energy/forms-of-energy.php (accessed Apr. 29, 2021).

[2] R. Ohlan, “Renewable and nonrenewable energy consumption and economic growth in India,” *Energy Sources, Part B Econ. Plan. Policy*, vol. 11, no. 11, pp. 1050–1054, 2016, doi: 10.1080/15567249.2016.1190801.

[3] “World energy consumption - Wikipedia.” https://en.wikipedia.org/wiki/World_energy_consumption (accessed Apr. 29, 2021).

[4] “EU Power Sector 2020 - Ember.” https://ember-climate.org/project/eu-power-sector-2020/ (accessed Apr. 30, 2021).

[5] T. O. Inventory *et al.*, “Estonia,” no. June 2020, 2022.

[6] H. K. Alfares and M. Nazeeruddin, “Electric load forecasting: Literature survey and classification of methods,” *Int. J. Syst. Sci.*, vol. 33, no. 1, pp. 23–34, 2002, doi: 10.1080/00207720110067421.

[7] X. Kong, C. Li, C. Wang, Y. Zhang, and J. Zhang, “Short-term electrical load forecasting based on error correction using dynamic mode decomposition,” *Appl. Energy*, vol. 261, no. December 2019, p. 114368, 2020, doi: 10.1016/j.apenergy.2019.114368.

[8] M. Spichakova, J. Belikov, N. Kalvi, and E. Petlenkov, “Feature Engineering for Short-Term Forecast of Energy Consumption,” *2019 IEEE PES Innov. Smart Grid Technol. Eur.*, vol. 00, pp. 1–5, 2019.

[9] “Estonia Population (2021) - Worldometer.” https://www.worldometers.info/world-population/estonia-population/ (accessed Apr. 11, 2021).

[10] “Choosing an electricity supplier | Eesti.ee,” Estonian government information portal | Eesti.ee. [Online]. Available: https://www.eesti.ee/en/housing-andenvironment/ services-related-to-housing/choosing-an-electricity-supplier/.” .

[11] A. Botterud, T. Kristiansen, M. I.-E. Economics, and undefined 2010, “The relationship between spot and futures prices in the Nord Pool electricity market,” *Elsevier,* Accessed: Apr. 01, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140988309002278?casa_token=NRTgHPKLFfoAAAAA:uvyO9Q02xrKd3fpO51ljhMrE4PC7KzX3x3bCjDy8k4QRD4-dKBIRhvNpd8X4giS6WqEZLyg9lPY.

[12] “Price of electricity to surge on Wednesday | News | ERR.” https://news.err.ee/1100084/price-of-electricity-to-surge-on-wednesday (accessed Apr. 01, 2021).

[13] “2020 warmest year on record in Estonia’s climate history | News | ERR.” https://news.err.ee/1608153526/2020-warmest-year-on-record-in-estonia-s-climate-history (accessed May 08, 2021).

[14] S. Karamizadeh, S. M. Abdullah, M. Halimi, J. Shayan, and M. J. Rajabi, “Advantage and drawback of support vector machine functionality,” *I4CT 2014 - 1st Int. Conf. Comput. Commun. Control Technol. Proc.,* no. I4ct, pp. 63–65, 2014, doi: 10.1109/I4CT.2014.6914146.

[15] “Weather in Tallinn (airport) - RP5.” https://rp5.ru/Weather_in_Tallinn_(airport) (accessed Apr. 11, 2021).

[16] “Nord Pool.” https://www.nordpoolgroup.com/ (accessed Apr. 05, 2021).

[17] G. A. N. Mbamalu and M. E. El-Hawary, “Vol. 8, No. 1,” vol. 8, no. 1, pp. 343–348, 1993.

[18] I. Moghram and S. Rahman, “Analysis and Evaluation of Five Short-Term Load

Forecasting Techniques," *IEEE Power Eng. Rev.*, vol. 9, no. 11, pp. 42–43, 1989, doi: 10.1109/MPER.1989.4310383.

[19] A. D. Papalexopoulos and T. C. Hesterberg, "5, No. 4, November 1990 1535," *Forecast*, pp. 1535–1550, 1990.

[20] Takeshi Haida and S. Muto, "Regression Based Peak Load Forecasting," *IEEE Trans. Power Syst.*, vol. 9, no. 4, pp. 1788–1794, 1994.

[21] T. Haida, S. Muto, Y. Takahashi, and Y. Ishi, "Peak load forecasting using multiple-year data with trend data processing techniques," *Electr. Eng. Japan (English Transl. Denki Gakkai Ronbunshi)*, vol. 124, no. 1, pp. 7–16, 1998, doi: 10.1002/(SICI)1520-6416(19980715)124:1<7::AID-EEJ2>3.0.CO;2-B.

[22] O. Hyde and P. F. Hodnett, "An adaptable automated procedure for short-term electricity load forecasting," *IEEE Power Eng. Rev.*, vol. 17, no. 2, pp. 48–49, 1997.

[23] M. Nazeeruddin, "Regression-Based Methodology for Daily Peak Load Forecasting," 1999, no. 1, pp. 468–472, [Online]. Available: http://library1.nida.ac.th/termpaper6/sd/2554/19755.pdf.

[24] S. A. A. R. E.H. BARAKAT, M.A. QAYYUM, M.N. HAMED, "On Power," *Prairie Schoon.*, vol. 5, no. 3, pp. 138–139, 1990, doi: 10.1353/psg.2016.0149.

[25] D. G. Infield and D. C. Hill, "Optimal smoothing for trend removal in short term electricity demand forecasting," *IEEE Power Eng. Rev.*, vol. 17, no. 12, p. 66, 1997.

[26] Q. C. Lu, W. M. Grady, M. M. Crawford, and G. M. Anderson, "An adaptive nonlinear predictor with orthogonal escalator structure for short-term - load forecasting," *IEEE Power Eng. Rev.*, vol. 9, no. 2, p. 49, 1989, doi: 10.1109/MPER.1989.4310472.

[27] M. M. Crawford and L. A. Groce, "Enhancement, Implementation, and Performance of an Adaptive Short-Term Load Forecasting Algorithm," *IEEE Trans. Power Syst.*, vol. 6, no. 4, pp. 1404–1410, 1991, doi: 10.1109/59.116982.

[28] j h P. Lee, K Y and S. Member, "COMPOSITE MODELING FOR ADAPTIVE SHORT-TERM LOAD FORECASTING," no. 2, pp. 450–457.

[29] J. Y. Fan and J. D. McDonald, "A real-time implementation of short-term load forecastjng for distribution power systems," *IEEE Trans. Power Syst.*, vol. 9, no. 2, pp. 988–994, 1994, doi: 10.1109/59.317646.

[30] Liu; S. Subbarayan; R.R. Shoults; M.T. Manry; C. Kwan; F.I. Lewis, "Comparison of Short-Term Load Forecasting Techniques," *2020 IEEE Conf. Technol. Sustain. SusTech 2020*, vol. 11, no. 2, pp. 877–882, 1996, doi: 10.1109/SusTech47890.2020.9150490.

[31] G. Juberias, R. Yunta, J. Garcia Moreno, and C. Mendivil, "New arima model for hourly load forecasting," *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.*, vol. 1, pp. 314–319, 1999, doi: 10.1109/tdc.1999.755371.

[32] H. Mori and H. Kobayashi, "Optimal fuzzy inference for short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 390–396, 1999, doi: 10.1109/59.486123.

[33] D. Srinivasan, C. S. Chang, and A. C. Liew, "Demand Forecasting Using Fuzzy Neural Computation, With Special Emphasis On Weekend And Public Holiday Forecasting," *IEEE Trans. Power Syst.*, vol. 10, no. 4, pp. 1897–1903, 1995, doi: 10.1109/59.476055.

[34] K. Padmakumari, K. P. Mohandas, and S. Thiruvengadam, "Long term distribution demand forecasting using neuro fuzzy computations," *Int. J. Electr. Power Energy Syst.*, vol. 21, no. 5, pp. 315–322, 1999, doi: 10.1016/S0142-0615(98)00056-8.

[35] N. Kandil, V. Sood, and M. Saad, "Use of ANNs for short-term load forecasting," *Can. Conf. Electr. Comput. Eng.*, vol. 2, pp. 1057–1061, 1999, doi: 10.1109/ccece.1999.808193.

[36] A. S. Joanicjusz Nazarko, "APPLICATION OF STATISTICAL AND NEURAL APPROACHES TO THE DAILY LOAD PIROFILES MODELLING IN POWER DISTRIBUTION SYSTEMS," *Japanese J. Crop Sci.*, vol. 45, no. 1, pp.

21–22, 1999.

[37] P. Lauret, M. David, and D. Calogine, “Nonlinear models for short-time load forecasting,” *Energy Procedia,* vol. 14, no. November 2015, pp. 1404–1409, 2012, doi: 10.1016/j.egypro.2011.12.1109.

[38] A. Rafati, M. Joorabian, and E. Mashhour, “An efficient hour-ahead electrical load forecasting method based on innovative features,” *Energy*, vol. 201, p. 117511, 2020, doi: 10.1016/j.energy.2020.117511.

[39] M. El-Hendawi and Z. Wang, “An ensemble method of full wavelet packet transform and neural network for short term electrical load forecasting,” *Electr. Power Syst. Res.*, vol. 182, no. February, p. 106265, 2020, doi: 10.1016/j.epsr.2020.106265.

[40] D. Howard and B. Mark, “Neural Network Toolbox Documentation,” *Neural Netw. Tool*, p. 846, 2004.

[41] “About website rp5.” https://rp5.ru/docs/about/en (accessed Apr. 12, 2021).

[42] I. A. Samuel, S. Ekundayo, A. Awelewa, T. E. Somefun, and A. Adewale, “Artificial neural network base short-term electricity load forecasting: A case study of a 132/33 kv transmission sub-station,” *Int. J. Energy Econ. Policy*, vol. 10, no. 2, pp. 200–205, 2020, doi: 10.32479/ijeep.8629.

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I Shan Ali Malik

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "Application of Artificial Neural Networks for Short Term Load Forecasting on Example of Publicly Available Estonian Data", supervised by Eduard Petlenkov.
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

10.05.2021

1 The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.