Available online at http://www.journalcra.com

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

International Journal of Current Research
Vol. 13, Issue, 03, pp.16506-16520, March, 2021

ISSN: 0975-833X

DOI: https://doi.org/10.24941/ijcr.40896.03.2021

REVIEW ARTICLE

OPEN ACCESS

USING VANILLA RECURRENT NEURAL NETWORK MODEL AND K-NEAREST NEIGHBOR ALGORITHM FOR PREDICT PRODUCTION AND CONSUMPTION ELECTRIC

***MOHAMMED ALQATQAT and Ma Tie Feng**

School of Statistics, Southwestern University of Finance and Economics, Chengdu, China

ARTICLE INFO

Article History:
Received 11th December, 2020
Received in revised form
27th January, 2021
Accepted 19th February, 2021
Published online 17th March, 2021

Key Words:

Electric, Temperature,
Learning Machine,
KNN algorithm,
Vanilla RNN Model.

ABSTRACT

Production and consumption electric prediction under the influence of temperature, interesting and challenging research topic Developed countries' economies are measured according to their power economy. Currently, the temperature is considered to be an illustrious scientist field because in many result it gives different. The temperature with its huge and dynamic information sources is considered as a suitable environment for data mining and business researchers. In this paper, we applied k-nearest neighbor algorithm and Vanilla RNN Model Suggested Method approach in order to predict Production and consumption electric prediction to assist investors, management, decision makers, and users in making correct and informed investments decisions. According to the results, the Vanilla RNN Model Suggested Method is robust with small error ratio; consequently the results were rational and also reasonable. In addition, depending on the Production and consumption electric data; the prediction results were close and almost parallel to actual Production and consumption electric prediction data.

Citation: MOHAMMED ALQATQAT and Ma Tie Feng. "Using vanilla recurrent neural network model and k-nearest neighbor algorithm for predict production and consumption electric", 2021. *International Journal of Current Research, 13, (03), 16506-16520.*

1. INTRODUCTION

The growing number of presumes on smart grids and the increasing contribution of renewable energy sources are due to increased energy availability and energy demand volatility. The development of complex forecasting systems that will replace traditional methods is a significant contribution to the sector. Some strategies require statistical methods based on historical data analysis with new strategies that can include broad data sets of various types, such as supply patterns, historical demand, and weather (1). Such applications include complex machine learning processes based on black-box models, which is a significant challenge for users with no technical knowledge. Therefore, they are widely used by extensive services or large network operators to improve temporary demand and supply speculation.

***Corresponding author:** ***MOHAMMED ALQATQAT*,**
School of Statistics, Southwestern University of Finance and Economics, Chengdu, China.

However, this is a common problem for small local power providers who do not have a wide range of data with various parameters required for complex models. For example, municipal resources still need a way to understand and apply such complex machine learning strategies. This paper provides an overview of the implementation of short-term energy demand design tools and forecasts for presume conditions in local energy systems. The revised version includes visual analytics and descriptive machine learning to simplify the analysis and prediction of power supply and Neighbor (kNN algorithm) for predicting power supply And demand produces a visual dashboard that enables users to analyse other predictive methods (2). Besides, it allows Users to understand how they relate to input Parameters. When identifying production and usage patterns.

2. SUMMARY OF LITERATURE REVIEW

The efficient management of the modern power system is a challenging problem in various countries across the globe. The smart grid's role is essential for handling different challenges in the power management system.

The smart grid assists in the energy management system's automation process by integrating various energy management systems such as computational ability, actuators, and sensors (3). The smart grid uses modern technologies to improve planning, production, operations, and maintenance. Across the globe, several governments sustain modern communication systems by keeping different energy and environmental saving factors. The smart grid carries out various tasks to simultaneously cope with different issues (4). According to (5), the US department of energy provides definitions of processes carried out by the smart grid in different ways. The energy suppliers mainly focus on urban power since the urban sector represents a high percentage of energy consumption. Various resources worldwide have been forced to focus on other aspects of power planning, management, demand, and supply (6). The increasing use of renewable energy sources such as wind power and solar energy and the active role of presumes that include the grid's remaining electricity have been forced to add more details to their electricity consumption and production analysis. Besides, presume has developed new strategies for successfully performing that analysis (7).

Various services worldwide provide dual business flow, including the purchase of generators and the provision of electricity. The concept of smart grids involves combining real-time energy production of various sources such as solar and wind power plants and identifying changes in energy demand and supply in real-time (8). In smaller services, traditional practices focus on heuristic predictions based on simple mathematical strategies or previous informing activities such as power supply and demand (9). However, this is not the case in the new context with the various influence factors and the increase in instability resulting from presumes' introduction into grids and renewable energy sources (10). Therefore, various resources worldwide are looking for new ways to get an accurate forecast of product productivity and efficiency. Power forecasting is a complex process that involves many dynamic identifications that creates energy demand; for example, the use of history and climate (11).

In many cases, the results are usually compiled into charts using different periods of yearly, monthly, or daily, and the information is seldom presented under the levels of use that they are finally able to understand. Linear Regression (LR) is a model that enables you to know the relationship between the return variable (other variables) and the variability of the response (power consumption) (12). The primary purpose of using regression analysis is because it is a method that triggers the prediction of energy demand from more than one cause (independent variable) (13), which can be alternating, housing availability, energy price, or day of the week. When the trend of historical predictive data is seen, then the linear regression method works very well. Due to its nature and simplicity, line retrieval has been used in various activities related to electricity use prediction. (14) Compared with a multi-country forecasting forecast, according to a complex economic model, such as Markel-Time, the improved retreat is consistent with official speculation. However, other studies, such as the one conducted by (15), studied the electrical prediction model in England. The model is based on several linear regression analyses, taking into account human and economic dynamics. (16) Researched various modelling methods to predict monthly energy consumption in Australia. These studies have shown outstanding results using mathematical models.

Support Vector Regression (SVR) has always been used to predict power consumption (17). Besides, SVRs have also been integrated with other methods to obtain better forecasting; for example, ant colony optimization that works to perform feature extraction while preventing training with large data sets that also do not work in some way (18), (19) Has developed an SVR-based process that enables the prediction of Singapore's energy use. In this case, the researcher first came up with a model for each variable, for example, exports, imports, and a complete country product, and these variables were combined to obtain predictable prices for use. The database is made up of one-year data, and the model was able to predict action over the next six years. One of the algorithms used for prediction is k-Nearest Neighbors (kNN) (20). Because of its simplicity, the K-algorithm is widely used, and it is easy to find similar conditions in the spaces of the larger and more diverse scales (21). However, this approach is limited when identifying previous causes of the same reliable variables to match future forecasts is not a predictable cause (22). Therefore, this method is mostly filled with temporary information such as a variant that identifies a week in the year, a day in the year, or a day of the week that helps search for similar Neighbors.

Temporary data entry is included in the data processing process (23). This method used has been used in various studies to make forecasts to predict the price of electricity and photovoltaic plants. Another model of study that is widely used to make predictions is Random Forests (RF) because it prepares a broad set of data that generates divisions with a high success rate (24). That is to say, kNN variables that gives a temporary value and should be used when talking about improving predictions made (25). One study that used RF with a high percentage of success in predicting electricity consumption in Argentina was conducted (26). The Gaussian Process Regression (GPR) is a powerful machine learning model used to create Bayesian trends in jobs. As it tends to be much better than other regression methods, the GRP model's reversal of the Gaussian process is much better than other regression methods, especially when access to adequate training data becomes problematic (27). Research by (28) shows that data collection is not too broad and how you can benefit from more than predictable results. A few researchers have made comparisons between different strategies such as Last Square SVM (LSSVM), Decision Tree (DT), Support Vector Machine (SVM), Extreme Learning Machine (ELM), and Autoregressive Integrated Moving Average (ARIMA). Comparisons of these methods revealed the success rate of GPR (29).

The use of machine learning techniques, in particular, predicts energy consumption and productivity is based on an extensive database with a variety of data types as factors that influence energy demand (30). Increased complexity has led to the development of predictive strategies based on complex algorithms, such as complex neural networks that are difficult to understand by mechanical and functional (31) learning specialists. As a result, the development of descriptive forecasting systems has, at the same time, demanded attention in line with the general movement of artificial intelligence (AI). The main goal is to improve the machine learning process and transparency to provide users with meaningful data regarding a given prediction model's performance (32). Depending on the (33), visibility can be achieved in three levels: the learning algorithm, individual items, and the entire model. As a result, simple machine learning algorithms, for example, kNN or decision trees are organized into a descriptive component (34). However, other strategies have examined the use of these methods in predicting energy demand. For example, kNN, they have not yet considered their explanatory power. Besides, they focused on checking the accuracy of various strategies to generate predictions for the next day automatically. The proposed method is very similar to the solution because we use the kNN algorithm to calculate the state's pre-sun power forecast in terms of minimum or maximum temperature as input parameters (35). Furthermore, different k-factor needs to be considered in this case. However, this approach is different not from claiming absolute weather parameters, for example, wind power and medium solar hours but primarily by looking at the algorithm effect for users in a way that enables them to understand the reasons for the predicted outcome and reach their conclusions. Visual and visual analytics to provide an indirect or indirect explanation of machine learning results is a useful method used by descriptive techniques. The whole machine learning process can be explained by singing visuals so that new users can easily understand. Visual acuity is often used to demonstrate alternative outcomes for different values of input parameters (36). Besides, it enables the user to analyse the correlations between the corresponding results and input parameters. Visual analytics techniques have been used extensively to improve decision-making processes in various domains, such as intelligent forecasting systems. The idea is derived from the analysis of comprehension models that focus on using analytics and observational data analysis to help users find novel-related information in a complicated situation (37). As noted by various studies, visual analytics techniques that provide presume with information about their production and application patterns positively impact individual decisions.

As a result, installing devices such as smart meters have enabled various resources to manage the growing number of data. Therefore, they need tools to gain insight from big data sets to support decision-making in power management activities, such as energy planning strategies. Visualize complex data in various ways that can interact, use, and test visual analytics systems to help users understand data relationships and patterns, thus gaining the necessary understanding to make informed decisions. The solution approach is based entirely on providing resource analysts with a tool that provides a concise overview of the effects of climate change and the forecast of electricity demand according to user-defined parameters (predictive nature — or predictable weather) (38). The prediction process is based on using a kNN algorithm to make the prediction model easier to understand. Therefore, this helps to identify very similar historical data dates based on parameters that include the date of forecasting. Therefore, keeping the input parameters and their link to the predicted result sound better, it is essential to limit the optimal set of weather data and the k-value that determines the number of similar days to be considered for the system (39). Besides, to support the analysis of the results and their understanding, the system recognizes both the predictions of the same dates and the same dates identified in the past (k-NN) as well as the possible production and use of electricity (4). It is essential to know that visuals are not static but presented in an interactive visual dashboard that uses analytics techniques to analyse different data and link it. In this way, a descriptive prediction process can help resource analysts understand the process and process they have collected. Using visual analytics provides a better understanding of the results and their implications in support of decision-making in power management (5).

From the current state of the art, ingenuity is evident, which is an excellent measure of machine learning models' success in various databases (6). The above review also revealed what changes are appropriate to include temporary information relevant to specific algorithms, such as RF or kNN. Energy forecasting is often classified into three main categories: short, medium, and long-term (40) (41). Firstly, it's important to notice that the forecasting lengths are directly defined with regards to the sample rate, the increment time between data entries (43). Hence, datasets with hourly or daily increments will define forecasting lengths differently as compared to minute increment datasets. For instance, a distribution company that supports revenue projection may define short term together to 5 years ahead and future as 20 years a head (42). Additionally, the operations group in an Independent System Operator may define short term as a couple of hours ahead, medium term as five days ahead, and future as fortnight a head (42).

The electricity consumption behaviour is inherently transient in nature, and therefore the consumption pattern (as elaborated later within the paper) will be formed by long-run dependencies. The thought of incorporating temporal dependencies of energy knowledge on completely different timescales employing a feature methodology was explored by Arahal et al. (43) and Fan et al. (44) within the context of short load prognostication. Recurrent neural networks (RNNs) square measure one family of algorithms which will accommodate dependencies between consecutive time steps. However, as mathematically shown by (Hochreiter and Schmidhuber), vanilla RNNs (i.e. those that don't account for long-run dependencies) suffer from the problem of vanishing/exploding gradient that makes learning long-run dependencies troublesome (45). Hochreiter and Schmidhuber (45) recommended continual neural networks with long short memory (LSTM) units as a doable answer to the vanishing gradient downside noticed in easy RNNs (45, 46). This allowed the RNN models with LSTM units to model each short and long run temporal dependencies in time-series knowledge. The deep RNNs provide a more expressive feature space, while accommodating both long and short-term temporal dependencies (as discussed by Arahal et al. (43)) using adaptive LSTM units. The sequence to sequence approach has been previously employed in speech recognition and MT applications (48, 49, 47), and while the approach has been previously utilized in short-term meteorology (50), its application in energy prediction context has been largely unexplored.

Energy consumption has been continuously increasing due to the rapid expansion of high-density cities, and growth in the industrial and commercial sectors. To combat negative environmental conditions, reduce operating costs, and identify energy savings opportunities, it is essential to efficiently manage energy consumption. Internet of Things (IoT) devices, such as widely used smart meters, are capable of measuring and communicating data about energy use; thus, they have created opportunities for improved energy management as well as for energy forecasting. Machine learning techniques build a mathematical model based on this historical data in order to make predictions or perform a different task. The common machine learning algorithms used for energy forecasting are not well-suited for time series problems. The electricity consumption behaviour is inherently transient in nature, and the consumption pattern (as detailed later in the paper) can be shaped by long-term dependencies. The idea of in-corpora ting temporal dependencies of energy data along di erent timescales using a feature method was explored by Arahal et al. (51) and Fan et al. (52) in the context of short-term load forecasting. Recurrent neural networks (RNNs) are one family of algorithms that can accommodate dependencies between consecutive time steps. However, as mathematically shown by Hochreiter and Schmidhu, vanilla RNNs (i.e. those which do not account for long-term dependencies) su er from the problem of vanishing/exploding gradient, which makes learning long-term dependencies di cult. As another approach for forecasting short-term gross annual electricity demand for Turkey, (Kucukali and Baris) applied symbolic logic. The proposed model used GDP because the sole independent parameter and captured the system behaviour of the amount 1970-2014 (53). In 2012, Bilgili et al. applied artificial neural network (ANN), rectilinear regression (LR), and nonlinear regression (NLR) to estimate the electricity consumption of the residential and industrial sectors in Turkey.

Installed capacity, gross electricity production, population and total subscribership were selected as independent variables. Prediction of the electricity consumption is predicated on two different scenarios and therefore the results of the three methods were compared (54). The comparisons showed good agreement between the actual data and forecasting results. Also, the performance values of the ANN method were better than performance values of the LR and NLR models

3. METHODOLOGY

3.1 K-Nearest Neighbor (KNN) Algorithm

Classification: The K-Nearest Neighbor (KNN) algorithm is a method to declassification of objects based on learning data that is closest to that object. Learning data is projected into a multi-dimensional space, where each dimension represents a feature of the data. This space is divided into sections based on the classification of learning data. A point in this space is designated class c if class c is the most common classification found in the k of the nearest Neighbors' of the point. Near or far Neighbors are usually calculated based on the Euclidean distance with a formula in below;

$$d_i = \sqrt{\sum_{i=1}^{p} (x_{2i} - x_{1i})^2} \tag{1}$$

X1= Sample Data
X2= Test / Testing data
i= Data Variable
d = Distance
p = Dimension Data

In the learning phase, this algorithm only stores feature vectors and classification of learning data. In the classification phase, the same features are calculated forest data (whose classification is unknown). The distance from the new vector to the whole the learning data vector is calculated, and the closest number of K pieces is drawn. That point only then is the classification predicted to be included in the largest classification of these points. The Best K value for this algorithm depends on the data. In general, a high K value will reduce the effect of noise on classifications, but create boundaries between each classification became more blurry. A good K value can be chosen by parameter optimization, for example by using cross-validation. Special cases where classification is predicted based on the closest learning data (in other words, K = 1) is called an algorithm nearest Neighbor. The accuracy of the KNN algorithm is greatly influenced by the presence or absence of features irrelevant, or if the feature's weight is not equivalent to its relevance to classification. The KNN algorithm has several advantages, namely robustness to training data which has a lot of noise and is effective when the training data is large. Meanwhile, weakness KNN is the KNN needs to determine the value of the K parameter (number of closest Neighbors), distance based training it is not clear what distance types to use and attributes which one should be used to get the best results, and the computation costs are quite high because it is necessary to calculate the distance from each query instance in the entire training sample. K-Nearest Neighbor (KNN) is a method that uses an algorithm supervised where the results of the new query instances are classified by majority from the category on the KNN. The purpose of this algorithm is to classify new objects based on attributes and training sample. The classifier does not use any models for matched and based only on memory. Given a query point, a number will

Table 1. Data Training

X1=DurabilityAcid (seconds)	X2 = Strength (Kg / sqm)	Classification
7	7	Bad
7	4	Bad
3	4	Good
1	4	Good

Be found object or (training point) closest to the query point. Classification of using the most votes among the classifications of K objects. The KNN algorithm uses classification Neighbor as the predictive value of the new query instance. As illustration of the application of the KNN algorithm, for example, there is result data survey. As the case, for example at this time the paper mill has produced a new network pass the laboratory test with X 1 = 3 and X 2 = 7. To guess the classification of this new network then calculations are performed using the KNN algorithm. The steps for calculating the K of the closest Neighbors is using the KNN algorithm are as follows.

Determine the parameter K (number of closest Neighbors). Suppose K = 3.

Calculate distance between query (data testing) and all practice examples (data training). The training data for which the proximity will be calculated has coordinates (3, 7), (Table 2).

Table 2. Distance Calculation

X1=DurabilityAcid (seconds)	X2 = Strength(Kg / sqm)	Square Distance to exampledemand (3, 7)
7	7	$(7-3)^2 + (7-7)^2 = 16$
7	4	$(7-3)^2 + (4-7)^2 = 25$
3	4	$(3-3)^2 + (4-7)^2 = 9$
1	4	$(1-3)^2 + (4-7)^2 = 13$

Table 3. Minimum Distance Rating Order

X1=Durability Acid (seconds)	X2=Power (Kg / sqm)	Square Distance to sample requests (3, 7)	Rating distance minimum	Is includingin-neighborsclosest 3?
7	7	$(7\text{-}3)^2 + (7\text{-}7)^2 = 16$	3	yes
7	4	$(7\text{-}3)^2 + (4\text{-}7)^2 = 25$	4	no
3	4	$(3\text{-}3)^2 + (4\text{-}7)^2 = 9$	1	yes

Table 4. Group Y Category nearest Neighbor

X1=Durability Acid (seconds)	X2=Power (Kg / sqm)	Square Distance to sample requests (3, 7)	Rating distance minimum	Is includingin-neighborsclosest 3?	Neighbors
7	7	$(7\text{-}3)^2 + (7\text{-}7)^2 = 16$	3	yes	Bad
7	4	$(7\text{-}3)^2 + (4\text{-}7)^2 = 25$	4	no	_
3	4	$(3\text{-}3)^2 + (4\text{-}7)^2 = 9$	1	yes	Good

Sort determine the nearest Neighbors distance and the closest distance-based to- C (Table 3)

Collect category Y from the closest Neighbor's row. On the second line the Neighboring category closest (Y) is not included because the data ranks more than 3 Neighbors closest (Table 4) above shows Group Y Category nearest Neighbor

Use the simple majority of the nearest Neighbor Category as the predictive value example query.

From the table above we get two new good quality tissue paper and one new tissue paper poor quality. Because the closest Neighbors who get more are of good quality, it can be concluded that the new tissue paper has passed the laboratory test with X 1 = 3 and 2 = 7 is included in the good category. The k-nearest Neighbor's algorithm is one of the strongest data classification techniques, by means of looking for cases by calculating the closeness between new cases and old cases based on matching weights (56). K-Nearest Neighbor is a supervised learning algorithm method, where the class that appears the most (majority) will be the class resulting from the classification. K-Nearest Neighbor is included in the instance-based learning group. K-nearest Neighbor is instance-based learning algorithm, where data sets training (training) is stored, so that the classification for new unclassified records is obtained by comparing records most similar to a training set. K-Nearest Neighbor steps:

- Determining the K parameter (the number of closest Neighbors), the K parameter in the test is determined based on the optimum K value during training
- Calculating the square of the Euclidean distance of each object against the sample data which is given.
- Sort the objects into groups that have the smallest Euclidian distance
- Collect category Y (nearest Neighbor classification).
- By using the majority category, the classification results can be obtained In general, to define the distance between two objects x and y, the distance formula is used Euclidian in equation.

Where the distance matrix is the scaled distance of both the x and y vectors of the matrix with size n dimensions. In the training phase, this algorithm only stores feature vectors and classification of training sample data. In the classification phase, the same features are calculated for testing data (whose classification is unknown), the distance of this new vector to all vectors.

Classification is the process of finding a model that describes and differentiates class's data, or how to classify data into one or more defined classes previous (57). Classification techniques are widely used, including Neural, Rough sets, K-nearest neighbor, Bayesian classifiers, network, and others. The data classification process consists of 2 steps, namely learning (training phase) and classification. The learning process is made to analyse training data and then it is represented in the form of a rule Classification. While the classification process, where test data is used to estimate accuracy of the classification rule. The model is built by analysing the tuple database. Every tuple is assumed to be a predefined class determined by an attribute called class label attribute (58). It can be illustrated in (Figure 1) below:

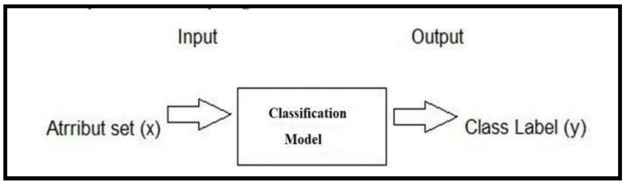

Fig 1. Model Classification

3.2. The classification process is based on four components:

- Class The dependent variable is a category that represents the "label" contained in the object.
- Predictor The independent variable represented by the characteristics of the data.

Advantage of using this model is that forecasting can be done more simply compared to the causal model.

Prediction: Prediction is the activity of predicting the values of a variable based on its values known from these variables or variables related. By their very nature, predictive techniques divided into 2 types, namely qualitative techniques and quantitative techniques. Grouped quantitative techniques into 2 types (59):

Time Series Model: In the time series model, future forecasting is based on future data values ago or called historical data. The purpose of this method is to find patterns in historical data series and utilize the sequence pattern for future forecasting. Time series data too describes an object over time or period historically and occurs sequentially. The advantage of using this model is that forecasting can be done more simply compared to the causal model.

Causal Model: The causal model is a model that assumes predicted factors show a causal relationship in one or more independent variables and use it to predict the future value of a dependent variable. An advantage in using this model is able to produce a greater success rate in retrieval wise decision.

Data Mining: Data mining is a term commonly used to describe findings knowledge in the database (60). Knowledge discovery data can process all non-trivial data to find patterns in the data, where the pattern is found to be valid and understandable.

The KDD stages are:

Data: Create a target data set, define a data set and focus on a subset variables or data samples, where the research will be carried out.

Data Selection: The first steps of data processing and data cleaning are basic actions such as noise removal. Before doing the data mining process, a cleaning process is required on data that is the focus in KDD

Transformation: At this stage is the stage of the creative process and is very dependent on information patterns will be searched in the database.

Data Mining: In selecting data mining algorithms to search for data mining processes, namely among others, techniques, methods or algorithms in data mining vary widely Determination method or the exact algorithm depending on the goals and overall KDD process.

Evaluation This stage is the stage of checking whether the patterns found contradict facts or a hypothesis that exists before, (Figure 2) shows the KDD process.

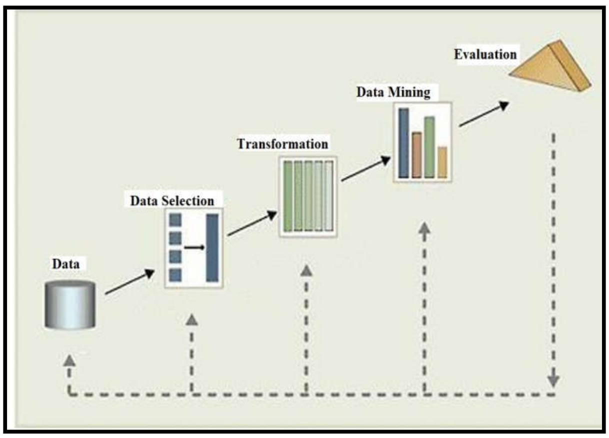

Fig.2. KDD process

Data Mining Techniques: The data mining function is to classify the patterns that must be found within data mining. The following are operations and techniques related to data mining (61):

- Predictive Modelling Operations: (classification, value prediction).
- Database segmentation: (demographic clustering, neural clustering).
- Link Analysis: (association discovery, sequential pattern discovery, similar time query discovery).
- Deviation detection: (statistics, visualization).

4. Vanilla RNN Model Suggested Method

4.1 Neural Network Models

- Models of neural static networks: such as multi-layered forward-fed neural networks, but without back feeding, where the outputs are calculated directly depending on their correlations with the front-feed inputs, and the response of this network at any point in time depends only on the value of the input sequence in The time stage itself, i.e. flowing in one direction from the inputs to the output.
- Models Networks Neural Dynamic: The output of this network depends
- On the current and previous values of the inputs and outputs or on the structure of the network where the response of this network is at any time A given that depends not only on the present value but on the past values of the input string. And dynamic neural networks can be classified into many networks, the most important of which are:

A- Neural networks - TDNN (Networks Neural Delay Time)
B- Neural networks with Windows Sliding with MLP
C- Networks Neural Recurrent: are neural networks with one or more feedback link, which can be of a general or private nature. Feedback allows the feedback networks to acquire accurate state representation, which makes it suitable for various dynamic applications such as (prediction or nonlinear modelling systems, etc.).

Application of multi-layered neural networks with mobile windows using the reverse propagation algorithm in data prediction: -

Dividing Time Series

- Group for training
- Group to check
- Set for testing

Estimating the Prediction Series Time-: The time series is a series of vectors or numerical values that depend on time, by applying the reverse propagation algorithm to the trained data set in the form of a mobile window to implement general education (Learning online,) The estimation of the time series is shown in the following points : -

Normalization data -: means preparing the data before processing it in relation to the model for artificial neural networks and the non-linear self-regression network model with an external variable for use in the neural network (61) training process. Therefore, the field of this data must fall within the limits of the activation function.

Determining the structure of the network: Topology network Determining-: The steps below determine the connection between the nerve cell, the number of hidden layers, and the number of neurons in each layer.

Determine the nodes of the input and how the neurons in the network are related to each other, the variables in the chain model

Temporal variables are in terms of the shifted variables, i.e.: -

$$Y_t = f\left(Y_{t-1}, Y_{t-2}, \ldots\ldots. Y_{t-p}\right) \qquad (2)$$

Since (p) represents the degree of self-regression, i.e. the regression of Yt on the previous values. It is difficult to determine the nodes of the inputs, as it represents a big problem towards the neural network designer. Therefore, several methods have been proposed, the most important of which are:

- Dependence on the significant self-correlation coefficients for the stable time series as mentioned in the self-regression model and the moving average.
- Dependence on the method of data collection in the case of a seasonal compound, for example: if the data contains an annual seasonality then (the number of entry nodes for the weekly data = 52 P, for the monthly P = 12, and for the season P = 4.

There is no ideal way to determine the number of neurons used in the hidden layer. Some researchers choose them with a number equal to or more temporally shifted variables, while others use statistical criteria, but it requires a very long period of time.

To determine nodes output, most researchers in the field of using neural networks in Prediction agreed that one prediction ahead step forward is sufficient for one node.

1. Network the Train
2. Determine the error criteria
3. Forecasting Forecast

The main objective of this stage is to calculate the future values of the time series that have been trained. In the case of forecasting one step ahead, it is by using the equation using the actual observation of all the variables shifted as an input to the network as shown: -

$$Y_{t+1} = f\left(Y_t, Y_{t-1}, Y_{t-2}, \ldots\ldots. Y_{t-p}\right) \qquad (3)$$

Where (p) represents the degree of self-regression over the previous values (number of shifted variables) (61)

4.2 Vanilla RNN Model

Repetitive Neural Network (Recursive NN, RNN) Repeatedly complex deep networks through tree, like neural network architecture. In essence, recurrent neural networks are an effective extension of a recurring neural network, and have different mathematical graphs. The recurrent neural network (recurrent neural network) and the recurrent neural network (Recurrent) are collectively called the recurrent neural network (RNN).

A repetitive neural network (RNN) belongs to a time repetitive neural network, and the relationship between neurons of a recurrent neural network forms a matrix; a repeating structural neural network belongs to a repeating structural neural network, and a repeating structural neural network uses symmetric neural networks.
Structural recursion builds more complex deep networks. Repetitive neural networks (RNNs) cannot handle the problem of exponential explosion or disappearance of weights with repetition, and it is difficult to capture long-term temporal correlations.

RNN is deep learning model which learn from sequences (62). RNN recursively takes the past output (y_{t-1}) and adds it to the current input (x_t). RNN current output (y_t) learns from the past sequence, using the y_{t-1}. (Figure.3) shows one RNN unit structure.

The sequence is passed on as an input to the current input. Therefore, y_t would be influenced by the past sequences. The past sequences carries on the past results with a combination. Therefore, sequential information effects all the outputs and carried on throughout a given sequence. presented the RNN's equation to demonstrate the recursion of the y_{t-1} to the x_t. W_r is the weight given to y_{t-1} and W_n is the weight given to the x_t. In our experiment, the RNN has 100 hidden units connected to each other sequentially to generate the best results.

$$y_t = tanh(W_r y_{t-1} + W_n\, x_t) \qquad (4)$$

The most basic machine learning program (MLP) is the Vanilla RNN. We will compute explicitly all the algorithm for this model and the rest of the discussion will rely more on the graphical representation and the intuition behind those models. Its representation is,

Where V, U, W are linear weighting vectors. The equations describing this model are,

$$\begin{aligned} a_t &= b + W s_{t-1} + U x_t \\ s_t &= \tan h(a_t) \\ o_t &= c + V s_t \\ p_t &= softmax(o_t) \end{aligned} \qquad (5)$$

The equation above is new layer which prepares the processed information for the output. In general, the inner layer dimensionality $\mathbf{s}_t$ will be huger since it is in charge of keeping the information from the past events and hence carry more information than that needed in the output.

4.3. Vanilla RNN Model suggested method

Ssuggested method forecasting consumption and production electric using vanilla RNN suggest to use vanilla RNN but instead of using the defult activation function which is (tanh) I suggest to use (elu)as activation function it seems that of all the recurrent neural networks that I tried the RNN with elu as activation performed the best in forecasting the electricity consumption and production from Jan 2010 to Dec 2019 and surprisingly it performed better than the state of art the LSTM and better than some new architectures like shifted RNN.

It's known that the vanilla RNN tends to suffer from vanishing or exploding gradients problems so you shouldn't use unbounded activation function like or elu in it and you should use a bounded activation function like tanh but surprisingly the elu activation function worked far more better than the tanh even when we scale down the inputs and the outputs to make easier on the tanh to process the data.

The model trained on the data from 2003 to 2009 and produced low MAPE error of the forecasted data from 2010 to 2019.

Ssuggested method steps:

- Load the xls file into python using panda's library.
- Convert the file to numpy and preprocessor the data.
- Create RNN network which takes 3 inputs and have one vanilla RNN layer contains 96 neurons with elu activation and dense layer contains one neuron as output layer.
- in every time step the network outputs its prediction about the consumption of the current month using this inputs the number of the current month(Jan=1, Feb=2), the consumption of the previous month, the temperature of the current month the pseudo code of the RNN model
- State = zeros array with the size of the number of neurons for each row in inputs previous state = state
- State = dot product (input weights, row) + dot product (state weights, state) return state

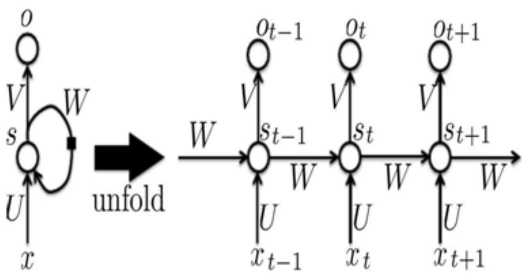

Fig.3.3. RNN unit structure

5. Data Description

The sample data was extracted from the https://www.ceicdata.com/en. The study sample included production and consumption electricity in china data and Average monthly temperature from https:// climateknowledgeportal.worldbank.org/ for 17 years, as a sample training dataset from the period Jan, 2010 to December, 2019 as shown in (Table 6, 7). Have approximately 120 records with three attributes including production electric, consumption electric, and temperature. A brief data analysis is presented with the fundamental concepts of data attributes. The attributes for each month are included in the data analysis. Production electric, consumption electric is the main factor that affects the prediction process for a specific temperature based on kNN Algorithm.

5.1 Model evaluation

The process of evaluating models is intended to evaluate the field suitability of the model for the pattern in which the series data is running or the accuracy of the model in predicting the values of the current and future series, and there are many measures of the suitability of the model all depend on the degree of error, which is the difference between the actual value of the series at a specific time And the string value that the model expected at that time. In this study, we will rely on the following methods to compare the two models used in this paper to find out which one is more accurate in prediction.

5.1.1 Mean Absolute Percentage Error (MAPE)

$$\text{MAPE}=100\sum_{i=1}^{n}\left[\left|Y_i-F_i\right|/Y_i\right]/n \qquad (6)$$

The mean absolute percentage error (MAPE), is a measure of prediction accuracy of a forecasting method in statistics, for example in trend estimation, also used as a Loss function for regression problems in Machine Learning. The MAPE (Mean Absolute Percent Error) measures the size of the error in percentage terms. It is calculated as the average of the unsigned percentage error

5. 1.2 Mean absolute deviation (MAD)

$$\text{MAD}=\frac{1}{n}\sum_{i=1}^{n}\left|x_i-\bar{x}\right| \qquad (7)$$

The median absolute deviation is a measure of statistical dispersion. Moreover, the MAD is a robust statistic, being more resilient to outliers in a data set than the standard deviation. In the standard deviation, the distances from the mean are squared, so large deviations are weighted more heavily, and thus outliers can heavily influence it.

6. RESULTS AND DISCUSSION

The result and analysis has been made using python (Waikato Environment for Knowledge Analysis), a Data mining specialized software. Python is a popular suite of machine learning written in Java, Developed by Guido van Rossum at Centrum Wiskunde & Informatics (CWI) in the Netherlands General Public It Python is a popular and general-purpose programming language.

Table 5. Group Y Category nearest Neighbor

1	x<1356,58	Very Very low
2	1356,58<x<2349,15	Very low
3	2349,15<x<3341,72	Low
4	3341,72<x<4334,3	Medium
5	4334,3<x<5326,87	High
6	5326,87<x	Very High

We Can write machine learning algorithms using Python, and it works well, the reason why Python is so popular among data scientists is that Python has a diverse variety of modules and libraries already implemented that make our life more comfortable. Our experiment uses the k nearest neighbor algorithm method and Vanilla RNN Model Suggested Method in order to predict the value of production and consumption electricity, based on a training data recorded from the period Jan, 2010 to December, 2019. The parameter temperature selected to create the different ranges for these Values are represented in (Table 5). This is a Drawback because the training set should have had more instances in order to create a model more accurate and precise. We start from our data in (Table 5) by calculating its mean, Standard deviation, Min and Max values please look then we define the states by their respective range.

Table 6- The results after applying kNN algorithm for the production electric

Model	Month	Jan		Fab		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
	year	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast
KNN Algorithm For Production Electric	2010	353	339	626	677.29	970	780	1308	1430	1657	1628	2009	1971	2399	2348	2795	2740	3144	3091	3485	3818	3830	3774	4200	4141
	2011	389	367	703	677	1091	1065	1468	1430	1855	1816	2252	2295	2687	3073	3124	3073	3516	3091	3895	3818	4284	4194	4703	4988
	2012	363	349	750	816	1166	1145	1555	1512	1962	1910	2376	2295	2833	2923	3283	3073	3688	3869	4088	3975	4503	4741	4976	4988
	2013	452	434	789	816	1214	1272	1630	1586	2057	2002	2496	2434	2990	2923	3500	2923	3945	3869	4383	4302	4831	4741	5420	5794
	2014	441	816	824	816	1279	1272	1715	1703	2164	2149	2628	2434	3137	3125	3640	3620	4098	4075	4548	4523	5012	4975	5638	5794
	2015	486	816	845	466	1290	1310	1732	1759	2189	2219	26629	2434	3167	3221	3678	3738	4134	4199	4584	4651	5049	5126	5802	5815
	2016	495	435	876	435	1352	1355	1809	1799	2282	2219	2776	2434	3329	3312	3892	3877	4389	4373	4878	4865	5385	5370	6130	5815
	2017	487	466	936	466	1446	1459	1931	1938	2426	2219	2951	4166	3558	3312	4157	4166	4689	4689	5202	4865	5733	5712	6364	6604
	2018	600	466	1055	523	1588	1459	2109	2220	2663	2636	3229	4166	3878	3312	4530	4166	5106	4689	5655	4865	6220	6163	6900	7112
	2019	617	545	1106	545	1680	1675	2233	2220	2799	2781	3398	4166	4065	3312	4742	4166	5344	5297	5923	4865	6514	6480	7226	7503

Table 7- The results after applying kNN algorithm for the consumption electric

Model	Month	Jan		Fab		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
	year	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast
KNN Algorithm For Consumption Electric	2010	353	339	626	677	969	779	1307	1430	1657	1627	2009	1970	2398	2347	2794	2740	3144	3090	3484	3818	3829	3774	4199	4141
	2011	388	367	703	677	1091	1065	1468	1430	1855	1816	2252	2295	2687	3073	3124	3073	3516	3091	3895	3818	4284	4194	4703	4988
	2012	363	349	750	816	1166	1145	1555	1512	1962	1910	2376	2295	2833	2923	3284	3073	3688	3869	4088	3975	4503	4741	4976	4988
	2013	452	434	789	816	1214	1272	1630	1586	2057	2002	2496	2434	2990	2923	3500	2923	3945	3869	4383	4302	4831	4741	5420	5795
	2014	441	816	824	816	1279	1272	1715	1703	2164	2149	2628	2434	3137	3125	3640	3620	4098	4075	4548	4523	5012	4975	5638	5794
	2015	486	816	845	466	1290	1310	1732	1759	2188	2219	2662	2434	3167	3220	3678	3738	4134	4199	4584	4651	5049	5126	5802	5815
	2016	495	435	876	435	1352	1355	1809	1799	2282	2219	2776	2434	3329	3312	3892	3877	4389	4373	4878	4865	5385	5370	6130	5815
	2017	487	466	936	466	1446	1459	1931	1938	2426	2219	2951	4166	3558	3312	4157	4166	4689	4689	5202	4865	5733	5712	6364	6604
	2018	600	466	1055	523	1588	1459	2109	2220	2663	2636	3229	4166	3878	3312	4530	4166	5106	4689	5655	4865	6220	6163	6900	7112
	2019	617	545	1106	545	1680	1675	2233	2220	2799	2781	3398	4166	4065	3312	4742	4166	5344	5297	5923	4865	6514	6480	7226	7503

Table 8- The prediction on test data production electric under Category kNN algorithm

Month	Jan			Fab			Mar			Apr			May			Jun			Jul			Aug			Sep			Oct			Nov			Dec		
year	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error
2010	1	1	0	1	1	0	1	1	0	2	2	0	2	1	1	2	2	0	3	3	0	3	4	-1	4	4	0	4	4	0	4	5	-1	5	5	0
2011	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	5	5	0
2012	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	1	2	4	4	0	4	5	-1	5	5	0	5	5	0	6	1	5
2013	1	1	0	1	1	0	1	2	-1	2	2	0	2	2	0	3	3	0	3	4	-1	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2014	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	1	3	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2015	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	4	-1	4	4	0	4	4	0	5	5	0	5	6	-1	6	6	0	6	6	0
2016	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	4	-1	4	4	0	5	4	1	5	5	0	6	6	0	6	6	0	6	6	0
2017	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	5	-2	4	1	3	5	5	0	5	6	-1	6	6	0	6	6	0	6	6	0
2018	1	1	0	1	1	0	2	2	0	2	3	-1	3	3	0	4	5	-1	5	4	1	5	5	0	6	6	0	6	6	0	6	6	0	6	6	0
2019	1	1	0	1	1	0	2	2	0	3	3	0	3	3	0	4	5	-1	5	4	1	5	5	0	6	6	0	6	6	0	6	6	0	6	6	0

Table 9- The prediction on test data consumption electric under Category kNN algorithm

Month	Jan			Fab			Mar			Apr			May			Jun			Jul			Aug			Sep			Oct			Nov			Dec		
year	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error
2010	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	4	4	0	5	5	0
2011	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	4	-1	4	4	0	4	4	0	5	4	1	5	5	0	5	6	-1
2012	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	5	0	5	5	0	5	6	-1	6	6	0
2013	1	1	0	1	1	0	1	2	-1	2	2	0	2	2	0	3	3	0	3	3	0	4	3	1	5	5	0	5	5	0	6	6	0	6	6	0
2014	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2015	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2016	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0	6	6	0
2017	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	5	-2	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0	6	6	0
2018	1	1	0	1	1	0	2	2	0	2	3	-1	3	3	0	4	5	-1	5	4	1	5	5	0	6	5	1	6	6	0	6	6	0	6	6	0
2019	1	1	0	1	1	0	2	2	0	3	3	0	3	3	0	4	5	-1	5	4	1	5	5	0	6	6	0	6	6	0	6	6	0	6	6	0

Table 10. The results after applying Vanilla RNN for the production electric

Model	Month	Jan		Fab		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
	year	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast
Vanilla RNN For Production Electric	2010	339	309	609	583	949	905	1284	1252	1628	1583	1971	1933	2348	2317	2740	2729	3091	3112	3422	3399	3774	3727	4141	4165
	2011	367	310	677	620	1065	1017	1430	1430	1816	1781	2217	2184	2643	2602	3073	3056	3454	3492	3818	3799	4194	4132	4604	4607
	2012	349	317	719	660	1145	1123	1512	1575	1910	1910	2295	2336	2744	2763	3191	3224	3583	3727	3975	3959	4384	4290	4988	4788
	2013	434	374	757	739	1182	1213	1586	1639	2002	1988	2434	2441	2923	2918	3432	3448	3869	3995	4302	4276	4741	4691	5432	5266
	2014	433	462	816	809	1272	1264	1703	1739	2149	2152	2616	2645	3125	3127	3620	3695	4075	4259	4523	4535	4975	4910	5794	5503
	2015	491	485	856	843	1310	1345	1759	1833	2219	2240	2709	2752	3221	3254	3738	3845	4199	4443	4651	4738	5126	5084	5815	5689
	2016	435	449	435	828	1355	1204	1799	1734	2268	2264	2759	2819	3312	3323	3877	3963	4373	4634	4865	4954	5370	5400	6133	6043
	2017	466	491	466	890	1459	1282	1938	1830	2437	2409	2960	3014	3570	3553	4166	4242	4689	4926	5194	5247	5712	5732	6604	6434
	2018	523	521	523	934	1576	1380	2088	1970	2636	2578	3195	3242	3837	3833	4480	4557	5036	5263	5582	5601	6163	6128	7112	6918
	2019	545	571	545	989	1675	1466	2220	2098	2781	2736	3367	3450	4030	4066	4703	4812	5297	5561	5874	5927	6480	6441	7503	7270

Table 11- The results after applying Vanilla RNN for the consumption electric

Model	Month	Jan		Fab		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
	year	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast
Vanilla RNN For Consumption Electric	2010	353	361	626	318	969	904	1307	1238	1657	1540	2009	1950	2398	2340	2794	2733	3144	3096	3484	3441	3829	3819	4199	4224
	2011	388	391	703	366	1091	1065	1468	1502	1855	1849	2252	2198	2687	2613	3124	3068	3516	3458	3895	3869	4284	4319	4703	4653
	2012	363	360	750	445	1166	1159	1555	1646	1962	2000	2376	2386	2833	2811	3284	3282	3688	3682	4088	4125	4503	4597	4976	4931
	2013	452	335	789	466	1214	1184	1630	1668	2057	2025	2496	2499	2990	2963	3500	3438	3945	3900	4383	4367	4831	4866	5420	5348
	2014	441	479	824	500	1279	1332	1715	1831	2164	2233	2628	2708	3137	3140	3640	3656	4098	4116	4548	4571	5012	5081	5638	5569
	2015	486	469	845	498	1290	1340	1732	1852	2188	2238	2662	2760	3167	3206	3678	3717	4134	4202	4584	4624	5049	5135	5802	5654
	2016	495	526	876	512	1352	1373	1809	1879	2282	2292	2776	2877	3329	3329	3892	3858	4389	4381	4878	4842	5385	5398	6130	6044
	2017	487	626	936	523	1446	1498	1931	2034	2426	2494	2951	3064	3558	3522	4157	4111	4689	4647	5202	5177	5733	5770	6364	6412
	2018	600	578	1055	541	1588	1541	2109	2142	2663	2620	3229	3223	3878	3801	4530	4418	5106	5019	5655	5609	6220	6248	6900	6905
	2019	617	625	1106	606	1680	1698	2233	2356	2799	2872	3398	3462	4065	4032	4742	4695	5344	5302	5923	5924	6514	6598	7226	7227

Table 12. The prediction on test data production electric under Category Vanilla RNN model

Month	Jan			Fab			Mar			Apr			May			Jun			Jul			Aug			Sep			Oct			Nov			Dec		
year	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error
2010	1	1	0	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	4	4	0
2011	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	4	4	0	5	5	0
2012	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0
2013	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	4	4	0	5	5	0	6	5	1
2014	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0
2015	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	5	-1	5	5	0	5	5	0	6	6	0
2016	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2017	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2018	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0	6	6	0
2019	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	4	4	0	4	4	0	5	5	0	5	6	-1	6	6	0	6	6	0	6	6	0

Table 13. The prediction on test data consumption electric under Category Vanilla RNN model

Month	Jan			Fab			Mar			Apr			May			Jun			Jul			Aug			Sep			Oct			Nov			Dec		
year	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error	Actual	Forecast	Error
2010	1	1	0	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	2	1	3	3	0	3	3	0	4	4	0	4	4	0	4	4	0
2011	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	4	4	0	5	5	-0
2012	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0
2013	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0
2014	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0
2015	1	1	0	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0
2016	1	1	0	1	1	0	1	2	-1	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2017	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0
2018	1	1	0	1	1	0	2	2	0	2	2	0	3	3	0	3	3	0	4	4	0	5	5	0	5	5	0	6	6	0	6	6	0	6	6	0
2019	1	1	0	1	1	0	2	2	0	2	3	-1	3	3	0	4	4	0	4	4	0	5	5	0	6	5	1	6	6	0	6	6	0	6	6	0

Evolutions / Model	MAPE Production electric	MAPE Consumption electric	MAD Production electric	MAD Consumption electric
KNN Algorithm Model	5.38%	10.51%	135.178	267.169
Vanilla RNN Model Suggested Method	5.93%	6.50%	86.404	83.144

7. CONCLUSION

The results of the predicted production and consumption electric used in the sample for the actual and predicted are presented. The results as seen in (Tables 6) and (Table 7) are those after applying kNN algorithm for production and consumption electric which indicates how far away is the predicted values from the actual values and similar for Vanilla RNN Model Suggested Method the result seen in (Tables 10) and (Table 11), The values of evaluation metrics per Model shown in (Table 14). The evaluation of the models for the consumption of electricity showed that Vanilla RNN Model Suggested Method performed best with MAPE (6.50%) and KNN Algorithm Model the MAPE (10.51%). The evaluation of the models for the consumption of electricity showed that Vanilla RNN Model Suggested Method performed best with MAD (83.144) and KNN Algorithm Model the MAD (267.169). The evaluation of the models for the production of electricity showed that Vanilla RNN Model Suggested Method performed best with MAPE (5.93%) and KNN Algorithm Model the MAPE (5.38%). The evaluation of the models for the consumption of electricity showed that Vanilla RNN Model Suggested Method performed best with MAD (86.404) and KNN Algorithm Model the MAD (135.178). The predictions on the test data are given in (Tables 8), (Table 9), (Table 12) and (Table 13) for each of the instances in the test set, the following is displayed: the Instance number, which is followed by the actual classification of the production and consumption electric, then the predicted classification, the last column represents the error classification: if the actual and the predicted values are equal, then the error is zero. Otherwise, it is displayed the error value: as a negative Number if the predicted value of the production and consumption electric is smaller than the actual one or as a positive number if the prediction gives a value greater than the actual value.

There are 106 of 120 instances with prediction error equal with zero, which represents a percent of 88.33% right predictions for KNN Algorithm Model. If a new data measurement is available, there it is a probability of 88.33% for this model to predict accurately the production electric. And there are 99 of 120 instances with prediction error equal with zero, which represents a percent of 82.5% right predictions for KNN Algorithm Model If a new data. Measurement is available; there it is a probability of 82.5% for this model to predict accurately the consumption electric. There are 115 of 120 instances with prediction error equal with zero, which represents a percent of 95.83% right predictions for Vanilla RNN Model Suggested Method. If a new data measurement is available, there it is a probability of 95.83% for this model to predict accurately the production electric. And there are 116 of 120 instances with prediction error equal with zero, which represents a percent of 96.6% right predictions for this model. If a new data measurement is available, there it is a probability of 96.6% for Vanilla RNN Model Suggested Method to predict accurately the consumption electric. This paper compared two models to forecast production and consumption electric energy for the period of 2003 to 2019. We find that the Vanilla RNN Model Suggested Method best and high accuracy more than KNN Algorithm Model in general after comparison. This work demonstrates the method of using Vanilla RNN Model Suggested Method for optimizing the predictions in the future for multiple effective characteristics. It can be concluded from the results and analysis that modelling using Vanilla RNN Model Suggested Method can be applied for forecast in others fields.

REFERENCES

Saini, I., Singh, D., & Khosla, A. QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *Journal of advanced research*, Vol. 4, Issue. 4, pp. 331–344, 2013

Saleh, A. I., Rabie, A. H., & Abo-Al-Ez, K. M.. A data mining based load forecasting strategy for smart electrical grids. *Advanced Engineering Informatics*, 30(3), 422-448. Vol. 30, Issue. 3, pp. 422–448, 2016

Samanta, S., Bera, J. N., & Sarkar, G. KNN based fault diagnosis system for induction motor. In 2016 *2nd International Conference on Control, Instrumentation, Energy & Communication* .India ,(pp. 304-308). 2016

Al-Faiz, M. Z., Ali, A. A., & Miry, A. H. (2010, November). A k-nearest neighbor based algorithm for human arm movements recognition using EMG signals. *Iraq J. Electrical and Electronic Engineering* Vol.6 Issue 2. pp. 158–166. 2010

Alkhalid, A., Lung, C. H., & Ajila, S. Software architecture decomposition using adaptive K-nearest neighbor algorithm. In 2013 26th IEEE *Canadian Conference on Electrical and Computer Engineering* Canada*, (*pp. 1-4). 2013

Arif, M., Malagore, I. A., & Afsar, F. A. Detection and localization of myocardial infarction using k-nearest neighbor classifier. *Journal of medical systems*, Vol. 36, Issue. 1, pp. 279–289, 2012

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm *machine learning approaches. Energies*, Vol. 11, Issue. 7, pp. 1636–1647, 2018

Serpen, G., & Aghaei, E. Host-based misuse intrusion detection using PCA feature extraction and kNN classification algorithms. *Intelligent Data Analysis*, Vol. 22, Issue. 5, pp. 1101–1114, 2018

Sharma, A., Jigyasu, R., Mathew, L., & Chatterji, S. Bearing Fault Diagnosis Using Weighted K-Nearest Neighbor. *International Conference on Trends in Electronics and Informatics* USA*, (*pp. 1321-1137). 2018.

Bryant, A., & Cios, K. RNN-DBSCAN: A density-based clustering algorithm using reverse nearest neighbor density estimates. IEEE Transactions on *Knowledge and Data Engineering*, Vol. 30, Issue. 6, pp. 1121–202, 2018

Danades, A., Pratama, D., Anggraini, D., & Anggriani, D. Comparison of accuracy level K-nearest neighbor algorithm and support vector machine algorithm in classification water quality status. *International Conference on System Engineering and Technology (ICSET).* Indonesia , (pp. 137-141). 2016.

Shen, X., Lin, Z., Peng, G., Tang, S., & Zhang, Y. Research on mechanical fault diagnosis of Ultra high voltage GIS based on the combination of neighbor algorithm and FCM. *Asia-Pacific Power and Energy Engineering Conference (APPEEC)* China, (pp. 2045-2049). 2016.

Song, J., Zhao, J., Dong, F., Zhao, J., Qian, Z., & Zhang, Q.. A novel regression modeling method for PMSLM structural design optimization using a distance-weighted KNN algorithm. *IEEE Transactions on Industry Applications*, Vol. 54, Issue. 5, pp. 4198–4206, 2018

Głowacz, A.. Diagnostics of induction motor based on analysis of acoustic signals with the application of eigenvector method and k-nearest neighbor classifier. *Archives of metallurgy and materials*, Vol. 57, Issue. 2, pp. 403–407, 2012

Glowacz, A. DC motor fault analysis with the use of acoustic signals, Coiflet wavelet transform, and K-nearest neighbor classifier. *Archives of Acoustics*, Vol. 40, Issue. 3, pp. 321–321, 2015

Güney, S., & Atasoy, A.. Multiclass classification of n-butanol concentrations with k-nearest neighbor algorithm and support vector machine in an electronic nose. *Sensors and Actuators B: Chemical*, Vol. 166, Issue. 2, pp. 721–725, 2012

Gupta, S., Reynolds, M. S., & Patel, S. N. Electric Sense: single-point sensing using EMI for electrical event detection and classification in the home. *In Proceedings of the 12th ACM international conference on Ubiquitous computing* USA, (pp. 139-148). 2010.

Souza, W. A., Marafão, F. P., Liberado, E. V., Simões, M. G., & Da Silva, L. C. A nil dataset for cognitive meters based on conservative power theory and pattern recognition techniques. *Journal of Control Automation and Electrical Systems*, Vol. 29, Issue. 6, pp. 742–755, 2018

Homaeinezhad, M. R., Atyabi, S. A., Tavakkoli, E., Toosi, H. N., Ghaffari, A., & Ebrahimpour, R. ECG arrhythmia recognition via a neuro-SVM–KNN hybrid classifier with virtual QRS image-based geometrical features. *Expert Systems with Applications*, Vol. 39, Issue. 2, pp. 2047–2058, 2012

Jakkula, V., & Cook, D. Outlier detection in smart environment structured power datasets. *International Conference on Intelligent Environments* .IEEE. Indonesia , (pp. 29-33). 2010.

Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K. C., & Hanzo, L. Machine learning paradigms for next-generation wireless networks. IEEE Wireless Communications, Vol. 24, Issue. 2, pp. 98–105, 2016

Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, Vol. 10, Issue. 1, pp. 841–851, 2017.

Lim, C. K. A., & Chia, W. C. (2015). Analysis of single-electrode EEG rhythms using MATLAB to elicit correlation with cognitive stress. *International Journal of Computer Theory and Engineering*, Vol. 7, Issue. 2, pp. 195–202, 2015

Lv, X., Cheng, X., & Tang, Y. M. Short-term power load forecasting based on balanced KNN. *In IOP Conference Series: Materials Science and Engineering* .Vol. 322, Issue. 7, pp. 072058, 2018

Tabatabaei, S. M., Dick, S., & Xu, W. Toward non-intrusive load monitoring via multi-label classification. *IEEE Transactions on Smart Grid*, Vol. 8, Issue. 1, pp. 26–40, 2016

Mahram, A., Shayesteh, M. G., & Jafarpour, S. Classification of wood surface defects with hybrid usage of statistical and textural features. *International Conference on Telecommunications and Signal Processing* .IEEE. Iran , (pp. 749-752). 2012.

Tian, J., Morillo, C., Azarian, M. H., & Pecht, M. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis. IEEE *Transactions on Industrial Electronics*, Vol. 63, Issue. 3, pp. 1793–1803, 2015

Majidpour, M., Qiu, C., Chu, P., Gadh, R., & Pota, H. R.. Fast prediction for sparse time series: Demand forecast of EV charging stations for cell phone applications. IEEE *Transactions on Industrial Informatics*, Vol. 11, Issue. 1, pp. 242–250, 2014

Martinez-Pabon, M., Eveleigh, T., & Tanju, B. (2017). Smart meter data analytics for optimal customer selection in demand response programs. *Energy Procedia*, 107, 49-59. Vol. 107, Issue. 4, pp. 49–59, 2017

Motlak, H. J., & Hakeem, S. I. Detection and classification of breast cancer based-on terahertz imaging technique using artificial neural network k-nearest neighbor algorithm. *International Journal of Applied Engineering Research*, 12(21), 10661-10668. Vol. 12, Issue. 21, pp. 10661–10668, 2017

Mustafa, M., Taib, M., Murat, Z., & Sulaiman, N. (2012). Comparison between KNN and ANN classification in brain balancing application via spectrogram image. *Journal of Computer Science & Computational Mathematics*, 2(4), 17-22. Vol. 2, Issue. 4, pp. 17–22, 2012

Trstenjak, B., Mikac, S., & Donko, D. KNN with TF-IDF based framework for text categorization. *Procedia Engineering*, 69, 1356-1364. Vol. 69, pp. 1356–1364, 2014

Nikravanshalmani, S. B. A. Hybrid algorithm based on K-nearest-neighbor algorithm and Adaboost with selection of feature by genetic algorithms for the diagnosis of diabetes, *International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)*. Vol 6. Issue 21, pp. 2977-2986. 2016

Qin, J., Wang, L., & Huang, R. Research on fault diagnosis method of spacecraft solar array based on f-KNN algorithm. In 2017 Prognostics and System Health Management Conference (PHM-Harbin) IEEE. China, (pp. 1-4). 2017.

Wahid, F., & Kim, D. (2016). A prediction approach for demand analysis of energy consumption using k-nearest neighbor in residential buildings. International Journal of Smart Home, Vol. 10, Issue. 2, pp. 97–108, 2016

Xie, Y., Deng, C., Liao, S., & Yuan, B.. Area-efficient K-Nearest Neighbor Design using Stochastic Computing. *Asilomar Conference on Signals, Systems, and Computers. IEEE*. USA, (pp. 782-786). 2018.

Yadav, A., & Swetapadma, A. Fault analysis in three phase transmission lines using k-nearest neighbor algorithm. *International Conference on Advances in Electronics Computers and Communications IEEE*. India, (pp. 1-5). 2014.

Zhang, C. Y., Chen, D., Yin, J., & Chen, L. Data-driven train operation models based on data mining and driving experience for the diesel-electric locomotive. *Advanced Engineering Informatics*, Vol. 30, Issue. 3, pp. 553–563, 2016

Ruiz, J. R. R., Canals, T., & Gomez, R. C. (2011). Comparative study of multivariate methods to identify paper finishes using infrared spectroscopy. IEEE *Transactions on Instrumentation and Measurement*.Vol.61, Issue.4, pp.1029-1036. 2011

Marino, D. L. K. Amarasinghe and M. Manic. Building energy load forecasting using deep neural networks. *Proceedings of the IEEE Industrial Electronics Society (IECON)*, Italy, *(*pp. 7046-7051). 2016.

Mocanu, E. P. H. Nguyen, M. Gibescu and W. L. Kling. Deep learning for estimating building energy consumption. Sustainable Energy, *Grids and Networks Journal*, Vol. 6. Issue .1: pp. 91–99, 2016.

Gupta A, Sarangi PK. Electrical load forecasting using genetic algorithm based back-propagation method. *Journal of Engineering and Applied Sciences*. Vol. 7. Issue .8: pp. 1017–1020, 2012

Arahal, M. R. A. Cepeda, E. F. Camacho, Input variable selection for forecasting models, *IFAC Proceedings* Vol.35. Issue .1: pp. 463–468, 2002

Fan, C. F. Xiao, Y. Zhao, A short-term building cooling load prediction method using deep learning algorithms, *Applied Energy*, Vol. 195©. pp. 222–233, 2017

S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation Vol.9. Issue .8: pp. 1735–1780, 1997

Lipton, Z. C. A critical review of recurrent neural networks for sequence learning, CoRR abs/1506.00019, http://arxiv.org/abs/1506.00019, Last Accessed = 2017-11-23.

K. Cho, B. Van Merrinboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, arxiv preprint arXiv:1406.1078. http://arxiv.org/abs/1406.1078

Sutskever, I. O. Vinyals, Q. V. Le, Sequence to sequence learning with neural networks, in: Advances in neural information processing systems, https://arxiv.org/abs/1409.3215

Zaytar, M. A. C. El Amrani, Sequence to sequence weather forecasting with long-short term memory recurent neural networks, *Internal Journal of Computer Applications*. Vol.143. Issue .11: pp. 7–11, 2016

Arahal, M. R. A. Cepeda, E. F. Camacho, Input variable selection for forecasting models, *IFAC Proceedings* Vol.35. Issue .1: pp. 463–468, 2002

Fan, C. F. Xiao, Y. Zhao, A short-term building cooling load prediction method using deep learning algorithms, Applied Energy Vol 195 (Supplement C) pp. 222 – 233. 2017

Kucukali S. and K. Baris, "Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach," *Energy Policy*, Vol.38. pp. 2438–2445, 2010

Bilgili, M. B. Sahin, A. Yasar and E. Simsek, "Electric energy demands of Turkey in residential and industrial sectors," *Renewable and Sustainable Energy Reviews*, Vol. 16, pp. 404-414, 2012.

Tanjung IH, 2012, " Forecasting the Total Sales of Egg Distributors on Demand Market Using Average-Based Fuzzy Time Series (ABFTS) *Journal of Statistical and Econometric Methods*, Medan Vol. 8, Issue .2. pp. 39-49, 2019.

MS Mustafa and IW Simpen, 2014 " Designing Appropriate Graduation Prediction Applications Time for New Students With Data Mining Engineering (Case Study: Academic Data STMIK Dipanegara Makassar Students) , " Citec J. , *Journal of Computer Application and Technology*,Vol.1. Issue .1: pp. 270–281, 2014

Mustakim, Giantika O, " The K-Nearest Neighbor Classification Algorithm as a System Prediction of Student Achievement Predicate , " J. Science and Techno. Ind. , Vol. 13, Issue. 2, pp. 195–202, 2016

Sumarlin, " Implementation of the K-Nearest Neighbor Algorithm as Decision Support Classification of PPA and BBM Scholarship Recipients , " J. Sist. Inf. Business, Vol. 01, pp. 52–62 ,2015

Fayyad, Usama, 1996. Advances in Knowledge Discovery and Data Mining. MIT Press.

Ian H. Witten, Frank Eibe, Mark A. Hall, Data Mining : Practical Machine Learning Tools andTechnique3rd Edition , Elsevier, Vol. 17, Issue. 3, pp. 37–57, 1996

Jiawei, H., Pei, J. and Kamber, M. Data mining: concepts and techniques. Elsevier, 2011.

Zhang, G. Peter, and Douglas M. Kline. (2007). Quarterly time-series forecasting with neural networks. IEEE transactions on neural networks ,Vol. 18, Issue. 6, pp. 1800–1814, 2007

Akintola, K. G., B. K. Alese, and A. F. Thompson. (2011). Time Series Forecasting with Neural Network: A Case Study of Stock Prices of Intercontinental Bank Nigeria. IJRRAS. Vol. 9, Issue. 3, pp. 467–472, 2011

Boukoros, S., Nugaliyadde, A., Marnerides, A., Vassilakis, C., Koutsakis, P., Wong, K.W.: Modelling server workloads for campus email traffic using recurrent neural networks. In: International Conference on Neural Information Processing, Greece. pp. 57-66. 2017

****