



Waleed Saleh Hamed *et al*, International Journal of Computer Science and Mobile Applications,
Vol.10 Issue. 10, October- 2022, pg. 1-15

ISSN: 2321-8363

Impact Factor: 6.308

(An Open Accessible, Fully Refereed and Peer Reviewed Journal)

Using Deep Learning Techniques to Predict Wind Speed

Waleed Saleh Hamed¹; Dr. Karim Q. Hussein²

¹Computer Science Department, College of Science, Mustansiriyah University, Baghdad, Iraq

waleedsalehhamed1979@gmail.com, ws987654321@yahoo.com

²Assistant Professor, Computer Science Department, College of Science, Mustansiriyah University, Baghdad, Iraq

karimzzm@yahoo.com, karim.q.h@uomustansiriyah.edu.iq

DOI: 10.5281/zenodo.7152925

Abstract: As a result of the rise in ecological pollution, the decrease in gas gets, and also the increase in power consumption against the rise in the population, researchers and scientists have motivated to concentrate their research on tidy as well as environmentally pleasant eco-friendly energy sources. Wind energy is a renewable resource that is cost-free from gas discharges during operation and does not need fuel costs. For that reason, wind power has actually drawn in those curious about producing in this work, the design was developed and also carried out based upon deep knowing methods for LSTM formulas, to execute several research studies to integrate this energy into the electrical power grid. SVR, KNN Regression on Kaggle World dataset. Models were examined utilizing analytical measures mean outright mistake (MAE) origin indicate square mistake, (RMSE). The version helps to enhance the forecast and reduce the issue of arbitrary modifications in wind rate.

Keywords— Wind speed prediction, Long Short-term Memory (LSTM), Deep learning (DL).





1. Introduction

The boost in environmental pollution with the boost in the worldwide population and the decrease in fuel books, encouraged researchers to focus their study to look for tidy and also eco-friendly energy sources, wind power is a extensive and sustainable energy source and does not create any kind of greenhouse gas exhausts when running, and also is taken into consideration a dispersant of toxins [1] Wind power generation has actually come to be a crucial component of electric power supply networks and also wise building networks, and due to the random characteristics as well as abrupt variations of the wind and also the low ability to properly predict the wind, it is an intermittent source of energy that has an effect on transmission and also distribution losses. Wind [2] Replacing thermal generation with wind power generation leads to conserving fuel expenses as well as reducing financial investment costs, as well as since wind power has no gas prices as well as does not produce greenhouse gases and also toxins, he motivated the combination of wind energy into electric power products [3]. Wind ranches are composed of large-capacity wind turbines, wind turbine systems can transform the power of the wind generators right into electric power, to raise power removal Wind generators are made for a high conversion rate, specific wind price projecting gives intending support for wind power generation as well as also choice of wind cattle ranch sites [4]. Specific projection can reduce the opposition between supply as well as need of power. Even more exact lots forecast is called for, as well as it is extremely vital to protect the operating safe and secure as well as steady of the grid system. As a result of the center non-linear consider between power tons and likewise non-linear facets, such as political setting, monetary strategy, human jobs and also irregular activity, it is actually tough to anticipate power tons exactly. Bunches projecting based upon historical information is the basic job of power system procedure, Therefore, to develop a much better projecting method finished up being additionally numerous as well as considerably essential scholars are committed to recommend a far much better prediction version to improve its forecast ability [5]. Long-term power sources are being considerably made use of in to provide greenhouse gas emission-free resources of electrical power in order to decrease the result on atmosphere adjustment. Wind power has really ended up being the globe's fastest broadening sustainable resource source of electrical power generation [6].





2. Related Work

Countless earlier research study has actually recommended many recommendations as well as additionally approaches. Despite of the methods made the most of, I acquired to manage boosting wind power predicting using deep understanding remedies, consisting of the following:

- **Arzu (2018) [7]** Using Of Multiple Linear Regression (MLR), Autoregressive Integrated Moving Average (ARIMA) together 'with Artificial Neural Networks (ANN) are the 3 reasonable strategies developed in this research study research. The style thinks that the well worth's are usually dispersed with a mean of no as well as in addition the distinction is continuous. The durability of the variation is figured out by screening assumptions. ARIMA variations call for the input data to have a continuous mean, alternative, as well as also furthermore autocorrelation utilizing time. in which we approximate the demands of feasible design(s) using the information at hand. 3 projection blunder actions are taken advantage of for style contrast along with analysis: Mean Absolute component Error (MAPE), Root Mean Square Error (RMSE) along with Mean Absolute Error (MAE). The outcomes revealed that the 3 versions sensibly forecasted Abasings wind cost in comparison to Persistence variation (standard) with ANN anticipating with a greater level of precision. The well worths of the varies to CNN were as comply with: MAE is 1.1863 along with RMSE 1.4822, MAPE 29.7312, R2 is 0.5505.
- **B.H. Mahdi (2021) [8]:** In this paper, ANN technique is recommended for forecast of WS in Duhok city, Iraq, utilizing Feed-Forward (FF) strategy for predicting DAWS. This research research study was done on the basis of the atmospheric info as inputs and wind rate as output for the duration of (2013-2018). consist of 2191 daily 'records for Duhok city, Iraq. Meteorological data that accumulated between 2013 to 2017 were utilized to educate the ANN, while the network was examined or had a look at taking advantage of 365 files (year 2018). to examine out the optimal format of the ANNFF, standing for the really finest choice coefficient, the least costly RMSE along with amongst one of the most inexpensive MAE The DAWS well worths are anticipated making use of 2 different conditions. The MAE in addition to RMSE well worths for forecasting DAWS were 0.98 as effectively as 1.34, particularly, for the first problem along with in enhancement 0.038 along with additionally 0.076, specifically, for the 2nd circumstances, It is clear to be observed that the recommended variant supplies extremely





specific end results. In enhancement, the MAE along with RMSE well worths program that the developed variant is certified of preparing yourself for future worths with high accuracy as contrasted to the exceptionally initial one.

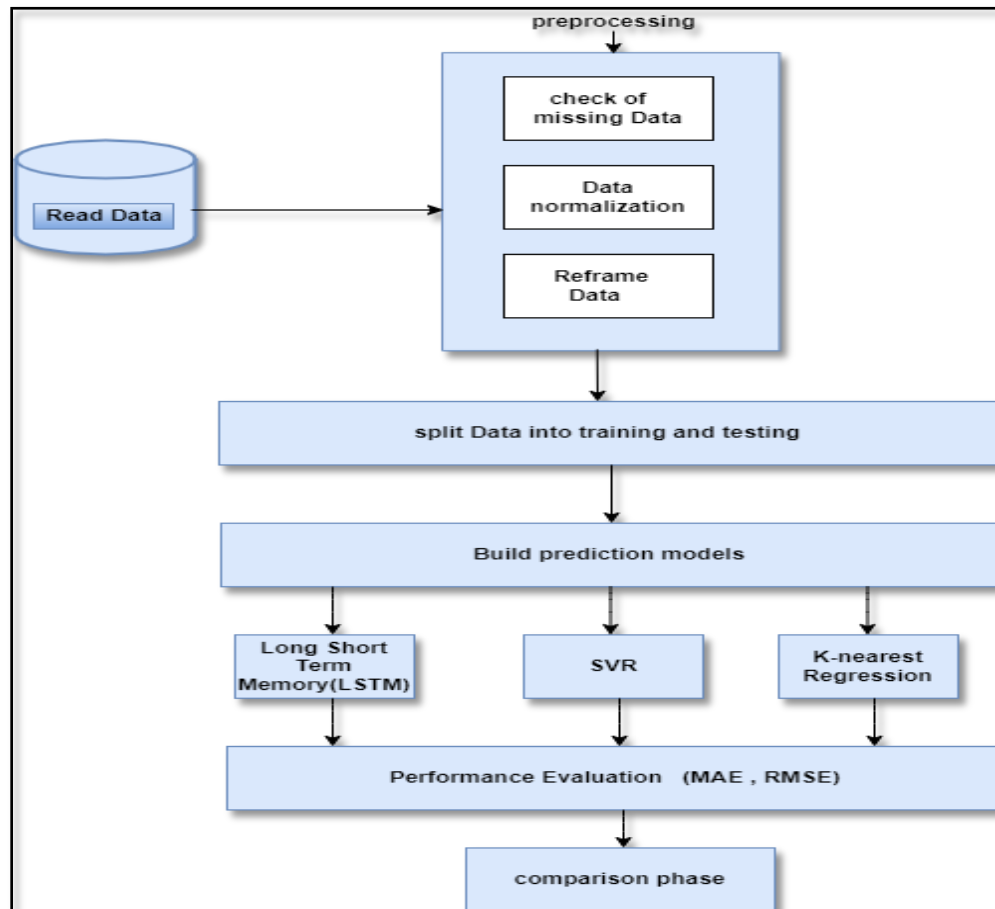
- **D. Jindong, (2022) [9]:** an one-of-a-kind crossbreed technique based upon wavelet modification (WT), DBN, and also PSO-DBN style in addition to spine quantile regression (QR) was advised to decrease the likewise nonlinear as well as non-stationary wind price collection. It generally concentrates on 3 critical elements. The extremely first is to create an included similarly precise as well as durable forecast format, Hence, LSTM, in addition to DBN, are picked as a low-frequency facet along with high-frequency part forecasters, specifically. The 2nd component is to boost the pertinent details preprocessing advancement to much better remove the initial details features. The 3rd part is the heavy mix of each sub-model. complete up that the nonlinear mix approach based upon PSO-DBN is a lot better than the linear mix technique. And likewise the preprocessing method successfully gets rid of the issue of the particular variant forecast hold-up.

3. Proposed Methods

In this paper, a prediction model was created as well as implemented making use of a formula (LSTM) for wind rate prediction was developed based upon 4 critical functions. The proposed solution in this research study was composed of 4 primary strategies as complies with:

- 1- Data preprocessing, Because valuable and vital details impacts the quality of the system as well as the rate of training, it is far better to pre-process the data before incorporating it into the system
- 2- Deep understanding approach is made use of to improve forecast outcomes forecast wind rate.
- 3- Evaluation treatment, we made usage of analytical steps to compute the efficiency of the prediction .Model.
- 4- contrast phase: for performance evaluation Where metrics are used mean outright mistake (MAE), and origin mean square error (RMSE). Figure 1 Describes the style of the system.





Fig(1) Schematic diagram of the proposed predictive model

3.1 Data Preprocessing

The raw data may include extreme values as a result of some equipment failing or practical errors, which might impact the accuracy of the version as well as the rate of training. The information have to be processed beforehand before training, as adheres to:

Step1: information cleansing.

Activity 2: Calculate the average for every row or column and replace the missing out on worth's with the standard for those attributes.



Activity 3: Normalization A normalization treatment for worths that have sound, which differs from the remainder of the worths because these values affect the prediction results, as well as make all values in between (0,1) or in between (1,-1) to low worths This is needed for great deals of algorithms.

Action 4: reformulating the information readied to anticipate the time collection information, the model should approximate the features from the previous worths & as a result, if the wanted output is $x(t)$ the input function should be $x(t-1)$. Collection to monitor function is created to take examples of $\{t-n, -n-1, ..., t-1\}$ as input for the machine learning. The output will be samples at $\{t, +1, ..., t+n\}$.

Action 5: Divide the data into a training and test set.

3.2 Deep Learning Models

The deep knowing technique is gotten its ability to solve varied troubles, and the capacity of DNN structures to enhance the forecast results of the deep knowing attribute, by including hidden layers in between the input and also outcome layers, as well as discovering algorithms are utilized experimentally to boost the training of features.

3.2.1 Long Short-Term Memory (LSTM): a deep understanding formula which has acceptable efficiency in prolonged short-term dependencies, is made use of to complete the predicting for the low-frequency sub-layers. [10] LSTM is made to get rid of these error heartburn problems. It can find out to connect time intervals in added of 1000 actions likewise in instance of loud, incompressible input series, without loss of short-time lag capacities. This is attained by a reputable, gradient-based algorithm for a layout using continual therefore, neither taking off neither going away error flow [11] LSTM networks have memory blocks that are connected with successive layers. Each block includes entries that take care of the state of the block and likewise the outcome. In the LSTM system. Expulsions bring weight that can be located out during the training stage. Certainly, expulsions make the LSTM obstructs smarter than timeless nerve cells and allow them to bear in mind current collection [12]. The first layer LSTM will certainly evaluate the details as they enter into the variation. Details that satisfies the regulations will be maintained in addition to details that does not accomplish the needs will certainly be neglected. Using this principle, the trouble of lengthy sequence reliance in semantic networks can be resolved. LSTM is an outstanding alternative of RNN. It inherits the benefits of



the RNN version as well as deals with the difficulty of the vanishing slope caused by the dynamic decrease of the gradient back breeding treatment there is a rounded chain in all flowing semantic networks. Each device of the chain structure improves the input details and also outputs it [13]. The LSTM system decreases or increases the transfer of info via 3 gateways containing: Forget, input in addition to result portals. This minimizes details. Every entry as received figure2

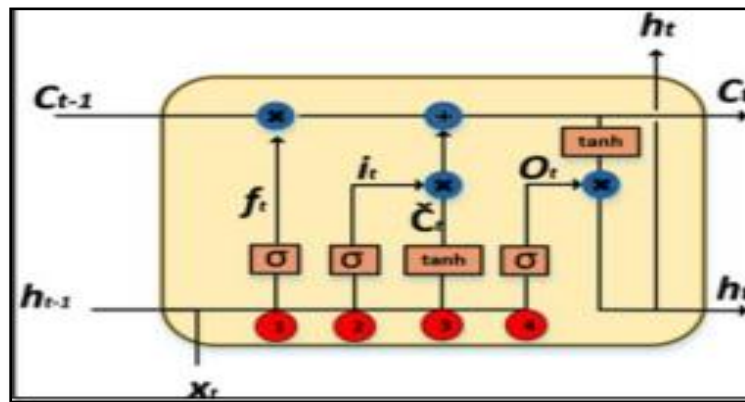


Figure (2) Demonstrates the LSTM structure

The very first entrance is the oblivion gateway whose feature is to define the details that will be disregarded in the previous memory cells, and it can be defined as adheres to [14].

$$f_t = \sigma(w_{hf} h_{t-1} + w_{xf} x_t + b_f) \quad (1)$$

The result of the neglect entry is increased by the old state to recognize how much old state details can enter the brand-new memory cell it can be specified as complies with [13].

$$\tilde{C}_t = f_t C_{t-1} + i_t \tanh(w_{hc} h_{t-1} + w_{xc} x_t + b_c) \quad (2)$$

Upgrading memory cells require usage overlook gateways as well as input gates according to the structure of LSTM. The possibility memory cell C_t is generated making use of the tanh attribute as information to go into the input gateway. can be defined as stick to

$$C_t = \tanh(w_{hc} h_{t-1} + w_{xc} x_t + b_c) \quad (3)$$

The second gateway is the input gateway, Its feature is to identify the amount of new details that can enter the memory cell C_t presently, and also it can be defined as adheres to:

$$i_t = \sigma(w_{hi} h_{t-1} + w_{xi} x_t + b_i) \quad (4)$$

The 3rd entrance is the outcome entrance, The surprise layer of the previous time action is updated by the result entryway which is also responsible for upgrading the output, it can 'be defined as complies with [14].

$$o_t = \sigma(w_{ho} h_{t-1} + w_{xo} x_t + b_o) \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

The Sigmoid feature is amongst one of one of the most made use of activation features as well as some times described as the logistic function. The sigmoid feature takes a real-valued number and transforms it into the range between 0 and also 1 which is specified as:.

$$\sigma(x) = \text{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (7)$$

The hyperbolic tangent (\tanh) function is an additional kind of activation function, the outcome variety of \tanh attribute between 1 along with -1 which is defined as:.

$$\sigma(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

the hyperbolic tangent function is more preferred in technique [15].

3.2.2 Support Vector Regression (SVR):. A support vector maker design (SVM) is made use of to fix the category issues of data in various courses, and the SVR design is made use of to resolve the forecast troubles. Regression is indicated to obtain a hyperplane that is fitted to the given data. The distance from any type of point on this hyperplane reveals the mistake of that certain point. The suitable strategy suggested for direct regression is the least-squares technique [16] SVR lowers the generalized error bound in contrast to decreasing the observed training mistake. To get the maximum generalization capability, the fundamental principle of SVR is to map the original info right into a high dimensional function area with nonlinear mapping [17].

3.2.3 K-nearest neighbors Regression (KNNR): It is an algorithm that can be utilized for classification 'and regression and also is made use of to approximate the thickness function for the circulation of test information and to categorize test data based on test patterns.

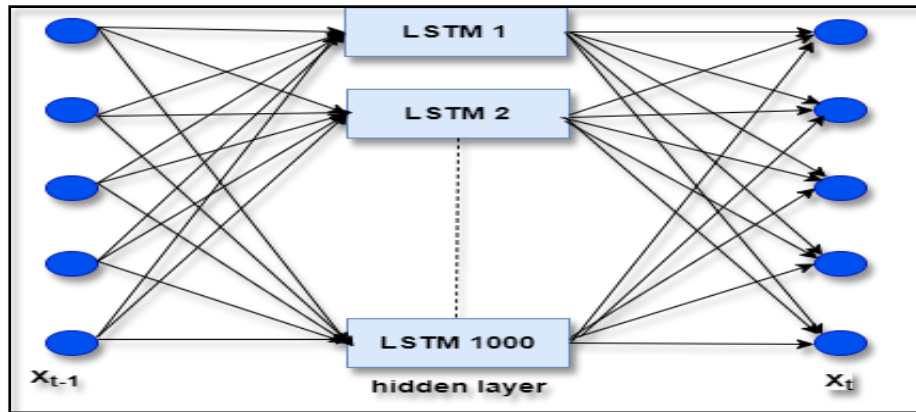


using this formula is to locate a method and also a relationship to compute the range in between the test and training information. Euclidean range is generally used to identify the distance in between examination and also training data After identifying the Euclidean range in between the data, the data source samples arranged in rising order from the least distance to the optimum distance to similarity. The next action Determining the variety of neighbors (k) is just one of the most important steps, and the performance of this technique is depended to the option of the closest (one of the most comparable) examples from the recommendation database substantially [16]. Prediction based upon KNN Regression can substantially lower training time cost savings and handling expenses, KNN has the ability to bypass the complex equation resolving procedure as well as concentrates on the correlation by utilizing raw data top qualities. The k-nearest neighbors selected from the training collection can establish the performance of projecting results. Consequently, the estimate of the new factor is connected to the values of its k-nearest neighbors. the KNN formula is associated to look for features of historic time collection that appear like the future electrical heaps. [18].

3.3 Building Models

In this paper, a forecast system was designed as well as carried out making use of LSTM formula based upon deep learning, with lasting reliance and also high efficiency to anticipate information with time-series as well as non-linear framework, and also being excellent with low-frequency wind layers, along with carrying out Support Vector Regression (SVR) and K algorithm -nearest neighbors Regression (KNNR) To contrast the efficiency of the results of the version, the design contains 5 fundamental input feature worths, a hidden LSTM layer including 1,000 nodes and an outcome layer of 5 info function values as shown in the complying with figure(3)





The figure (3) shows the LSTM model

4. Dataset Description

Wind speed data set collected from Kaggle World website for the year 2012 for each measured time hour. This data set contains 16,800 record and 6 columns. The dataset shows in Table

Table (1): Sample of a dataset

1	TIMESTAMP	TARGET VAR	U10	V10	U100	V100
2	20120101 1:00	0.000000	2.124600	2.681976	2.864280	3.666086
3	20120101 2:00	0.054879	2.521695	-1.796960	3.344859	2.464761
4	20120101 3:00	0.110234	2.672210	-0.822526	3.508448	1.214093
5	20120101 4:00	0.165116	2.457504	-0.143642	3.215233	0.355556
6	20120101 5:00	0.156940	2.245898	0.389576	2.957678	0.332701



For forecast of time series information, Features must be estimated from previous worths So, if the preferred outcome is $x(t)$, the input attribute ought to be $x(t-1)$ Series to monitored function is designed to take samples of $\{t-n, t-n-1, \dots, t-1\}$ as input for the deep learning. The output will be samples at $\{t+1, \dots, t+n\}$. and therefore, the input is the previous day and the result is the existing day. Table 2 showed the results of reframing data

Table 2: reframe date results

	var1(t-1)	var2(t-1)	var3(t-1)	var4(t-1)	var5(t-1)	var1(t)	var2(t)	var3(t)	var4(t)	var5(t)
1	-1.047630	0.465135	0.303671	-0.832710	-0.618776	-0.858000	0.620467	0.416316	-0.529624	-0.379867
2	-0.858000	0.620467	0.416316	-0.529624	-0.379867	-0.666725	0.679344	0.454661	-0.195909	-0.131143
3	-0.666725	0.679344	0.454661	-0.195909	-0.131143	-0.477085	0.595357	0.385933	0.036584	0.039598
4	-0.477085	0.595357	0.385933	0.036584	0.039598	-0.505336	0.512583	0.325564	0.219194	0.176472
5	-0.505336	0.512583	0.325564	0.219194	0.176472	-0.464420	0.410933	0.254713	0.358485	0.285841

In this design, we divided the dataset into 80% training and 20% testing. For optimization, we reformulate the information into a 3D matrix, time intervals, and features as input to the model.

5. Results and Discussions

In order to review the prediction and the aspects impacting the forecast results, we made a set of experiments based on LSTM innovation, which is an efficient formula for information of time collection, long-range dependency as well as low-frequency wind layers. Include 1000 nodes as well as finally an outcome layer that includes five nodes representing time t , and also the loss values for the version remained in epochs 50 as well as patch dimension 2500 for training and also verification details as received the number (4)



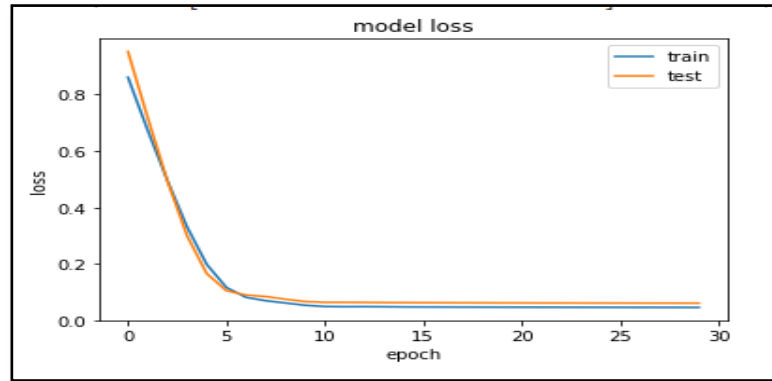
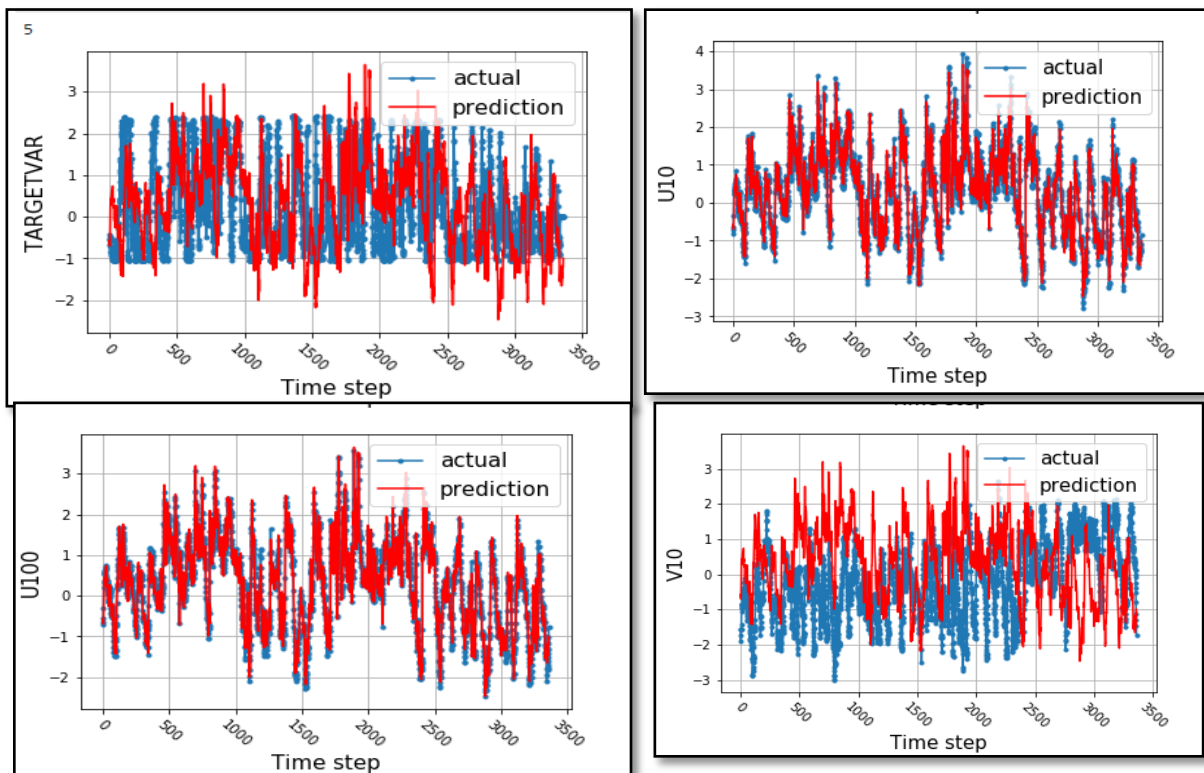


Figure (4) LSTM model loss function

In epoch 30, the training loss is equal to (0.0466) and validation loss is equal to (0.0617).

5.1 Experience and implementing all the features in (LSTM):

The model was trained on all features and the prediction results were as in Figure (5).



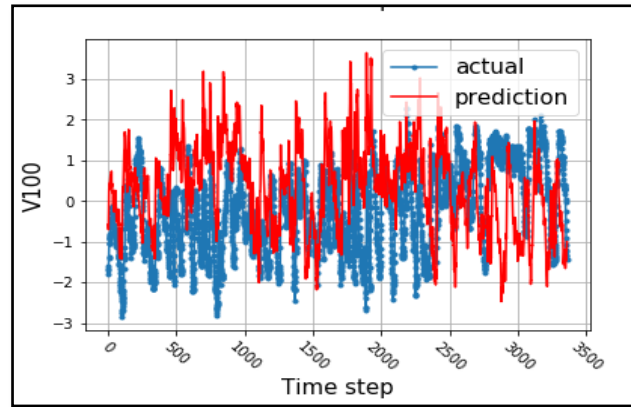


Figure (5) Demonstrates the accuracy of all features

Making use of the LSTM model to its capacity to predict long-lasting and not shed the initial attributes that were made with the reformulation of the elements, to gauge the performance as well as accuracy of the forecast by, imply absolute mistake (MAE) in the formula (9) as well as origin suggest square mistake (RMSE) in the equation (10).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_t - \hat{y}_t| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_t - \hat{y}_t)^2} \quad (10)$$

where \hat{y}_t and y_t represent the forecast value and actual value at time t , N denotes the complete number of the information [9]. According to the mentioned criteria, we conclude that the values of the LSTM model for MAE are (0.168), and the RMSE is (0.235).

SVR, KNN Regression and LSTM designs are examined to find the very best result among them by comparing the outcomes acquired from the accuracy dimensions MAE as well as RMSE, see Table (3).



Table 3: MAE and RMSE for prediction models.

Prediction models	predicting model errors	
	MAE	RMSE
(LSTM)	0.168	0.235
SVR	0.899	1.121
KNN Regression	0.866	1.057

observed from Table 3, the amount of the MAE as well as RMSE values of SVR as well as KNN Regression and LSTM so that the LSTM is superior to the various other techniques.

CONCLUSIONS

A prediction version based upon deep discovering techniques was studied and also created using numerous experiments on the information establish collected from Kaggle World on both LSTM, SVR, KNN Regression. Via the outcomes of the table (), the superiority of the LSTM version over the various other versions appears, as well as the model reveals a reliable efficiency for anticipating wind rate. Through the outcomes of the design, it is possible to predict the wind speed in the long term and to benefit from it in the installation of electric stations and the supply of energy networks.

REFERENCES

- [1] Mahdi, BH., Yousif, KM, Melhum, A. I. (2021).” Application of Artificial Neural Network to Predict Wind Speed: Case Study in Duhok City, Iraq”, In *Journal of Physics: Conference Series* (Vol. 1829, No.1, p. 012002).
- [2] Wang Y, Zou R., Liu F., Zhang L, Liu Q. (2021).”A review of wind speed and wind power forecasting with deep neural networks:”, *Applied Energy*, 304, 117766.
- [3] Mujeeb S., Alghamdi TA., Ullah S., Fatima A., Javaid N., Saba T.(2019).”Exploiting deep learning for wind power forecasting based on big data analytics”, *Applied Sciences*, 9(20), 4417.
- [4] Liu, H., Mi, X., & Li, Y. (2018). Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Conversion and Management*, 159, 54-64.
- [5] Nie, Y., Jiang, P., & Zhang, H. (2020). A novel hybrid model based on combined preprocessing method and advanced optimization algorithm for power load forecasting. *Applied Soft Computing*, 97, 106809.





- [6] Li, G., & Shi, J. (2010). On comparing three artificial neural networks for wind speed forecasting. *Applied Energy*, 87(7), 2313-2320.
- [7] Arzu, A., Kutty, S. S., Ahmed, M. R., & Khan, M. G. (2018, December). Wind speed forecasting using regression, time series and neural network models: a case study of Kiribati. In *Australasian Fluid Mechanics Conference*.
- [8] Mahdi, B. H., Yousif, K. M., & Melhum, A. I. (2021, March). Application of Artificial Neural Network to Predict Wind Speed: Case Study in Duhok City, Iraq. In *Journal of Physics: Conference Series* (Vol. 1829, No. 1, p. 012002). IOP Publishing.
- [9] Duan, J., Wang, P., Ma, W., Fang, S., & Hou, Z. (2022). A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting. *International Journal of Electrical Power & Energy Systems*, 134, 107452.
- [10] Liu, H., Mi, X., & Li, Y. (2018). Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Conversion and Management*, 159, 54-64.
- [11] Hochreiter, S., & Schmid Huber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [12] Abdel-Nasser M., Mahmoud K.(2017). "Accurate photovoltaic power forecasting models using deep LSTM-RNN". *Neural Computing and Applications*, 31(7), 2727-2740.
- [13] Wang, B., Kong, W., Guan, H., & Xiong, N. N. (2019). Air quality forecasting based on gated recurrent long short term memory model in Internet of Things. *IEEE Access*, 7, 69524-69534.
- [14] Abass, N. S., Sahib, N. M., & Al-Hayat, A. (2021, September). Forecasting of Individuals' Movement in Society After Corona Pandemic Using RNNs with LSTM and GRUs. In *2021 4th International Iraqi Conference on Engineering Technology and Their Applications (IICETA)* (pp. 117-122). IEEE.
- [15] Rysbek, D. (2019). *Sentiment analysis with recurrent neural networks on Turkish reviews domain* (Master's thesis, Middle East Technical University).
- [16] Shabani, S., Samadianfard, S., Sattari, M. T., Mosavi, A., Shamshirband, S., Kmet, T., & Várkonyi-Kóczy, A. R. (2020). Modeling pan evaporation using Gaussian process regression K-nearest neighbors' random forest and support vector machines; comparative analysis. *Atmosphere*, 11(1), 66.
- [17] Wang, J., Zhou, Q., Jiang, H., & Hou, R. (2015). Short-term wind speed forecasting using support vector regression optimized by cuckoo optimization algorithm. *Mathematical Problems in Engineering*, 2015.
- [18] Dong, Y., Ma, X., & Fu, T. (2021). Electrical load forecasting: A deep learning approach based on K-nearest neighbors. *Applied Soft Computing*, 99, 106900.

