Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, No. 2, No.01 (2023),
ISSN : **1906-9685**

EXPLORING ADVANCED DEEP LEARNING TECHNIQUES FOR RELIABLE WEATHER FORECASTING

Dr.Jitendra Singh Chouhan Head Department of Computer Applications Geetanjali Institute of Techincal Studies
Ms.Anisha Sheikh Assistant Professor Department of Computer Applications Geetanjali Institute of Techincal Studies
Mr.Ankur Rao Assistant Professor Department of Computer Applications Geetanjali Institute of Techincal
Mr.Ajay Jain Assistant Professor Department of Computer Science, B. N. University, Udaipur

***Abstract*—** Weather forecasting has always been a domain where more and more reliability is looked up to when it comes to prediction and the models pursued. While in the past years, traditional methods included conventional probability-dominated dependency models, the reliability of these traditional methods was always in a contentious line which emphasized the necessity to improve and use modern technology. While in the past few years, we became familiar with the ways that can put the huge amount of data to an optimum utility, it only widened the horizons which makes room for us to explore various important domains where reliability is indeed a necessity, that could very much use the help of advanced technology to be more decisive. The materializing deep learning techniques tied up to the humongous availability of observed data concerning the field of weather prediction have enticed enough researchers to traverse the concealed hierarchical pattern in the voluminous weather datasets. The study was done to inquire into various deep learning techniques including discussing some of the advanced and recently under-the-light methods.

Keywords*—*Artificial Neural Network, Long Short Term Memory, Rainfall, Weather, Naïve Bayes, SVM

I. INTRODUCTION

Worldwide meteorological department is putting great efforts into research zones of weather prediction. Since India is a horticulture-based country the vast majority of individuals are subject to the climate conditions. A huge amount of the Indian populace relies on monsoons. It is always a depend of a challenge for scientists and researchers to precisely forecast climate conditions. The weather forecast includes a prediction of rain, mist, fog, winds, clouds, lightning, hurricanes, etc. One of the most important challenges in weather forecasting is its unpredictable and dynamic weather information units, which could frequently trade consistent with global climatic situations.

The weather situation at any instance can be represented by way of a few capabilities. Out of these functions, one determined that the most significant is being selected to be concerned with the manner of prediction [1]. The selection of features is dependent on the vicinity for which the prediction is to be made. The capabilities and their variety usually vary from vicinity to location. The climate situation of any day has dates dating with the weather situation existing inside the equal tenure of treasured 12 months and previous week. Rainfall is the shape of precipitation. Its correct forecasts caidentify future viable floods and and to the rationale for higher water control. Weather forecasts may be categorized as Forecasts which forecast as much as a few hours, Short period forecasts are specifically Rainfall forecasts is 1 to three days forecasts, Forecasts for 4 to 10 days are Medium-range forecasts, and Long period forecasts for extra than 10 days[2]. Short-variety and Medium Range rainfall forecasts are significant for flood forecasting and water resource control.

II. BACKGROUND STUDY

The concept of weather forecasting using deep learning can be defined in a simplified manner as a method with a particular approach in which, based on the previously observed pattern or behavior of weather over a given area and timeline, considering n number of factors as well as their dependencies, the weather conditions of the future time is predicted [3]. More emphasis is laid upon the reliability of the models used by testing them frequently along with the help of testing datasets and once the model proves itself reliable enough based on the accuracy measure, the model is set ahead to be trained with the training dataset.

This considerable amount of data can be requested from any credible source or repository. For multiple and crucial applications, weather prediction is quite a necessary domain including military utilities, agriculture, production, aviation industry, etc these are some of the verticals where weather forecasting plays a major role in mitigating risk.

Precise weather prediction is a tactful task because of the dynamic nature of the earth's atmosphere, at any instance the weather condition over a particular demographic can be represented by some variables. Some of these variables are more significant and hence decisive than others and hence with the help of redundancy techniques priority can be assigned to these variables [4].

The above-discussed background information laid the grounds for us to move ahead. exploring research papers and whitepapers [5].

Given the wide scope of methods, one has to choose which modal serves the purpose best as per the requirement. Some of the algorithms and concepts that were frequently mentioned in the concerned research papers are-

- Classification under the supervised learning techniques is a model to differentiate samples with unknown magnificence labels on the premise of similarities and dissimilarities and predict a category label for them.
- The Naive Bayes approach and K-Medoids are the approaches that are quite frequently commended to determine the weather.
- The Bayesian technique for the category is a statistical and linear classifier that predicts elegance labels for facts example on the idea of the distribution of characteristic values. This is a parametric type wherein the dimensions of the classifier remain constant.
- The k-medoids algorithm is a clustering algorithm associated with the k-way algorithm and the k-medoid shift set of rules.

III. LITERATURE REVIEW

This paper [6] investigates deep studying strategies for climate forecasting. In specific, this observation compared the prediction performance of Recurrence Neural Network (RNN), Convolutional Network (CN) models, and Conditional Restricted Boltzmann Machine (CRBM). Recurrent neural networks can forecast rainfall with a high enough degree of accuracy. Future time-collection concerns may be predicted correctly using Deep Learning approaches like Conditional Restricted Boltzmann Machine (CRBM) and Convolutional Neural Network (CNN), which give the right representation, type, and prediction. Model accuracy is assessed using the Forbenius Form.

In this paper [7], the use of Naïve Bayes and K- medoid algorithm has been used for weather prediction systems with features such as temperature, humidity, and wind. The prediction can be considered reliable because the forecast would be based on previous records.

The predicting of the highest and lowest temperatures for a period of seven days was the exclusive focus of this article [8]. Both a functional regression model variant and a linear regression model were applied. Functional regression turned shown to be a high-bias, low-variance model while linear regression was found to be a low-bias, high-variance model. Adding extra data can help the linear regression model since it is an inherently high-variance model because it is unstable to outliers. Functional regression, which cannot be enhanced by more data.

The primary goal of the study is to provide an overview of the various artificial neural network (ANN) and data mining weather forecasting methods. Precipitation, temperature, and wind speed are the three variables that are most frequently used to analyse the [9] weather prediction. According to the study, using ANN, fuzzy logic, and data mining techniques leads to greater accuracy.

In this paper [10], they've focused on a new Python API for amassing weather records from websites. For all of the experiments, the facts turned into divided into 3independent sets: 60% for training, 20% for hyperparameter tuning (validation), and 20% for testing the accuracy prediction of models. Data was tested on two different types of ANN models i.e. an AR-NN and an ARX-NN model. The ARX-NN model included precipitation, whereas the AR-NN model simply took temperature as an input. The performance of the ARX-NN model was marginally better than that of the standard AR-NN model.

This [11] investigates three ML models for climate prediction, particularly ANN a Time Series (primarily based on RNN), and SVM. RNN is the usage of time series along with a 5-layered neural network and a linear SVC and is used to forecast the climate. Based on the Root Mean Squared Error (RMSE) between the actual values and the projected values, the model outputs are examined and contrasted. After analysis of all the models, it is observed that time series RNN is better than SVM and ANN for this Problem. Time Series RNN had RMS values of 1.41, ANN had a value of 3.1, and SVM had a value of 6.67.

This paper [12] proposes a rainfall prediction model with the use of Multiple Linear Regression (MLR) along with Linear Regression(LR). The proposed technique uses Indian meteorological statistics to predict rainfall. 70% of the records are for training and 30% of the records are for testing. The accuracy, correlation, and Mean Square Error (MSE), are the parameters used to verify the model proposed. From the above results, the MLR (MSE: 11.894, RMSE: 3.449, Correlation: 0.473) model provides better results than LR(MSE: 13.28, RMSE: 3.644, Correlation: 0.469).

The backbone of the Indian economy is agriculture. The monsoon is what farmers rely on for their agriculture. The ideal climate, fertilizer, and soil conditions are necessary for the best crop productivity. Forecasting the weather is a crucial need for any farmer. The meteorologist center [13] provided the data needed for this assignment. The case information spanned the years 2012 through 2015. At this stage of the study, the following tactics have been used: Data mining, data transformation, data cleaning, and data selectionThe most effective method, which produces precise and quick results, is ANN. It excels at delivering superior results for multidimensional data, boasts great processing speed, delivers increased precision, and successfully handles complicated real and imaginary values.

It is well known that numerical climate prediction (NWP) models need powerful computers to solve difficult mathematical equations to produce a forecast that is mostly dependent on current climatic conditions. It proposes a novel lightweight information-driven climate forecasting version by examining the temporal modelling approaches of long short-term memory (LSTM) and Temporal Convolutional Networks (TCN). Evaluating its performance of the current classical the suggested deep model consists of a few layers that forecast climate using floor climate parameters throughout a certain period. The proposed LSTM and TCN layers deep learning networks are evaluated in various regressions, especially a multi-input single-output and a multi-input multi-output [14].

For meteorologists, accurately and reliably predicting wind speed is challenging. The researchers used a variety of nonparametric tree-based analytic approaches along with specific convective weather characteristics to estimate the maximum wind speed at 10 metres. 127 convective storms that happened between 2005 and 2013 were the basis for the research. For the evaluation of point estimates and prediction intervals, the study evaluated the Bayesian Additive Regression Trees (BART) and Quantile Regression Forests (QRF) error models [15]. Several metrics that took into account the prediction intervals as well as the bias and random error of point estimates were used to evaluate the accuracy of the error models.

IV. OBJECTIVE

The problem with the existing weather prediction is that it is based on numerical weather prediction which requires more computation power to solve the mathematical equations therefore we need a model or system which is more reliable, accurate, and has a greater speed. As India is an agriculture-based country and it is also the primary source of income and revenue. The majority of the population is dependent on it so now it is more important to predict accurate weather not only it will help in agriculture but also for a better plan for water resource management, flood management, military, scientific research, and many more [16].

The objectives for this model are:

- To predict the weather more accurately and have a greater speed.
- The model should predict the weather for 12h or more and can be easily deployed and portable.
- Consider all the important features for prediction.

V. Theoretical Aspects

A. *Linear Regression:*

Regression problems are the main focus of the machine learning method known as linear regression, which is widely used in supervised learning. To create a predictive model for the target variable, it makes use of independent variables. In order to determine the link between variables and predictions, regression analysis is frequently used. The type of intercorrelation between the input and output elements taken into account, as well as the inclusion of the independent variables, can cause regression models to vary [17].

On the basis of a known independent variable (x), linear regression is used in this study to predict a structured variable price (y). As a result, the linear connection between x (input) and y (output) is shown by this regression approach.

B. *Cost Function*

The version aims to predict y fee such that the error differential between expected price and real value is minimised by attaining the exceptional-suit regression line. Therefore, updating the 1 and 2 values is absolutely necessary to arrive at a great value that reduces the error between predicted y cost (pred) and actual y cost (y) [18].

The Root Mean Squared Error (RMSE) between the real y value (y) and the predicted y charge (pred) is the cost function(J) of linear regression.

C. *Gradient Descent*

The version uses gradient descent to swap out $\theta 1$ and $\theta 2$ variables in order to minimise the Cost function (minimising RMSE charge) and arrive at the great match line. The concept starts with random values for $\theta 1$ and $\theta 2$, which are then updated repeatedly until the minimal value is reached [19].

D. *LSTM*

A supervised Deep Neural Network known as the Long Short Term Memory (LSTM) model excels in time-series prediction. It resembles a recurrent neural network in several ways. An LSTM version seems at closing “n” days (time step) statistics (additionally referred to as lag) and predicts how the collection can progress in the future. A artificial recurrent neural network (RNN) architecture used in deep learning is called LSTM. LSTM has feedback connections as opposed to traditional feedforward neural networks [20]. It cannot use the basic method of separate statistical elements (like photos), but rather entire statistical sequences (together with audio or video). For instance, LSTM may be used for tasks like detecting anomalies in network visits or IDSs, as well as unsegmented, connected handwriting popularity, voice popularity, and other similar tasks. A cell, an input gate, a forget gate, and an output gate make up a common LSTM unit. The three gates change the amount of records entering and leaving the cell, and the cell recalls data over arbitrary time intervals.

However, there are some editions of the LSTM version which includes Gated Recurrent Unit (GRUs) that do not have the output gate. LSTM Networks are popularly used on time-collection records for typing, processing, and making predictions [21]. The purpose of its recognition in time-collection application is that there can be several lags of unknown periods between important events in a time collection.

E. ANN

Artificial Neural Networks additionally alluded to as ANN is a scientific model dependent on organic neural networks. The artificial neural network is dependent on how the human mind perceives. The human brain is a very complex network of interconnected neurons [22]. ANN is essentially utilized in coordinating examples and making forecasts on a given arrangement of data sources.

VI. DESIGN ISSUES

The design issues in ANN are:

A. Initial Weights

These are small random values ranging from -1 to 1.

B. Transfer Function

These functions define how to combine inputs and weight to p; produce the output. Some of the transfer functions are:

i. Linear:

The yield is corresponding to the complete weighted info.

ii. Threshold:

The yield is set at one of two qualities, contingent upon whether the complete weighted information is more prominent than or not exactly some edge esteem.

iii. Non-Linear:

The yield fluctuates continuously yet not directly as the info changes.

C. Error Estimation

A common measure for measuring the discrepancy between observed or estimated values and the anticipated values produced by a model or estimator is root mean squared error (RMSE).

D. Weights adjusting

After every cycle, loads ought to be changed to limit the blunder. This is finished by every conceivable weight and backpropagation.

i. Backpropagation

At each layer, controlled learning is being used in this instance to reduce the discrepancy between Iayer's reaction and the actual information [23]. Each concealed layer's error is the average of all assessed errors. This is how hidden layer networks are taught.

- N is a neuron.
- Now is one of N’s inputs weights
- N0ut is N’s output.
- Nw = Nw +Δ Nw
- Δ Nw = N0ut * (1- N0ut)* NErr0rFactor
- NErr0rFact0r = NExpected0utput – NActual0utput

This works just for the last layer, as we can know the real yield and the normal yield

E. No. of neurons

The ANN model additionally relies upon the number of neurons.

i. If there are many neurons then the accuracy is high but the model becomes slow and there is a risk of over-fitting.
ii. If there are few neurons then the accuracy is lower and the model is unable to learn properly.
iii. Then comes the optimal number of neurons.

F. Data Representation

Normally the info information and yield information need pre-preparing for pictures we need pixel force and for content, we need an example [24].

G. Size of the training set

No preparing set fits all recipe. Overfitting can happen if a decent preparation set isn't picked. In a decent preparation set an example must speak to an inclusive community, must contain individuals from each class, and each class must contain a wide scope of varieties [25]. The size of the preparation set is identified with the quantity of concealed neurons.

H. ANN Architectures

Neural Networks can approximate all the functions present in the universe. Many methods are present for the approximation of all the nonlinear functions present in the universe. As with many types of architecture, the same variety of complexity is faced [26].

- Feedforward Networks

They have whatever number of shrouded layers as could be expected under the circumstances. The systems with just one shrouded layer have ended up being the most effective ones. There is no input inside the system [27]. The coupling takes royal residence starting with one layer and then onto the next. The data streams, all in all, are in a forward course.

i. Input layer: The number of contributions to the neural system is equivalent to the number of neurons in this layer. This layer has a few hubs that don't do anything. They simply give the data starting with one neuron and then onto the next

ii. Hidden layer: In this layer, we have any number of layers and any number of neurons. The decision of picking the number of neurons and layers is kept comp, lately irregular. The hubs in this layer are said to be dynamic since they are a piece of sign adjustment.

iii. Output layer: this layer gives us the quantity of yield which is equivalent to the number of neurons present in this layer. Every one of the hubs present participates in change.

- Feedback networks

The yield of the neuron present after the shrouded layer is straightforwardly or in a roundabout way feed into the neurons that are of late connected.

- Lateral networks

This layer implies coupling neurons in a specific layer. The yield of any neuron isn't nourished into some other neuron. This system is something in the middle of the feedforward system and criticism arrangement.

VII. DESIGN & IMPLEMENTATION

Till now we have seen the information gathered from the various papers we studied and the theoretical aspects of the main concepts related to the paper. In this paper, we will discuss the design we opted for to build our model and how we implemented our idea.

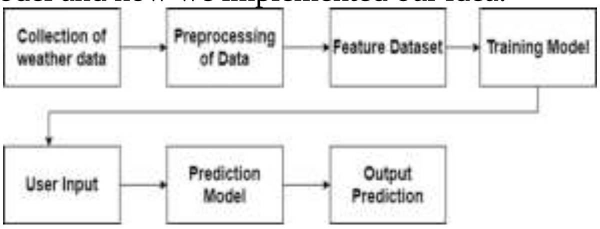

Fig. 1. Architectural Design of the work

Over a two-year period, from 4 May 2016 to 11 March 2018, a weather forecasting model was trained. The dataset is particular to the Indian city of Jaipur.

Discuss Input / Output requirements

The obtained data contains 40 features and 676 instances:

- Date
- meantempm
- maxtempm
- mintempm
- meantempm
- minpressurem
- precipm
- precipm

Fig. 2. Dataset

Experimental Scenarios:

- We will perform this experiment using different machine learning approaches like Linear Regression using the OLS method and Artificial Neural Network.
- We will compare the results or predictions obtained by different algorithms and find the best suitable one which can accurately predict the values.

VIII. RESULT & ANALYSIS

A. *Linear Regression with the OLS method*

We started with retrieving the dataset which is particularly based on Jaipur. The dataset has records from 4 May 2016 to March 11, 2018, followed by pre-processing of the dataset. Then we read some research papers to find out the best possible solution to the problem statement leading to which we decided to use Linear Regression with the OLS method and Artificial Neural Network.

One approach to survey the linearity between autonomous variables, which until further notice will be the mean temperature, and the other free factors is to compute the Pearson connection coefficient. To evaluate the connection in this information we will call the corr() strategy for the Pandas DataFrame object. This will yield the relationship esteems from the most adversely connected to the most emphatically associated.

```
In [48]: df.corr()[['meantempm']].sort_values('meantempm')
Out[48]:
```

	meantempm
minpressurem_1	-0.830438
minpressurem_2	-0.809432
minpressurem_3	-0.795442
meanpressurem_1	-0.794987
meanpressurem_2	-0.776774
maxpressurem_1	-0.764999
meanpressurem_3	-0.764715
maxpressurem_2	-0.745532
maxpressurem_3	-0.732154
maxhumidity_1	-0.254148
maxhumidity_2	-0.232007
maxhumidity_3	-0.214156
precipm_1	0.032026
precipm_2	0.048588
precipm_3	0.077058
minhumidity_1	0.085576
minhumidity_2	0.092412
minhumidity_3	0.094188
mindewptm_3	0.382905
mindewptm_2	0.400483
mindewptm_1	0.412313

Fig. 3. Pearson Coefficient

Features having less absolute value than 0.6 will be removed from the dataset. We have seen that the maxtempm and mintempm variables are useless as for the prediction of meantempm we will eliminate them likewise and supplant the nan esteems utilizing the fillna() strategy.

Fig. 4. Selecting Features

We will use the matplotlib pyplot module to graph the linear relationship between the independent variable and the dependent variable.

```
%matplotlib inline

# manually set the parameters of the figure to and appropriate size
plt.rcParams['figure.figsize'] = [16, 22]

# call subplots specifying the grid structure we desire and that
# the y axes should be shared
fig, axes = plt.subplots(nrows=6, ncols=3, sharey=True)

# Since it would be nice to loop through the features in to build this plot
# let us rearrange our data into a 2D array of 6 rows and 3 columns
arr = np.array(predictors).reshape(6, 3)

# use enumerate to loop over the arr 2D array of rows and columns
# and create scatter plots of each meantempm vs each feature
for row, col_arr in enumerate(arr):
    for col, feature in enumerate(col_arr):
        axes[row, col].scatter(df2[feature], df2['meantempm'])
        if col == 0:
            axes[row, col].set(xlabel=feature, ylabel='meantempm')
        else:
            axes[row, col].set(xlabel=feature)
plt.show()
```

Fig. 5. Plotting of Graph

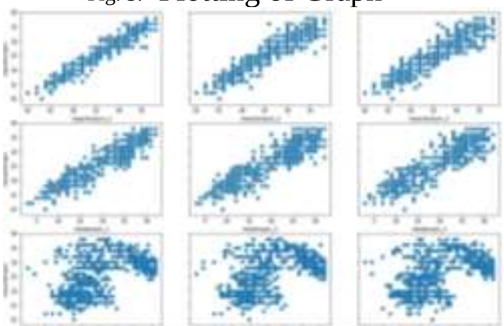

Fig. 6. Linear Relationship between Variables

To perform Linear Regression on the dataset, we have used the OLS algorithm. As we know that a linear regression model utilizes statistical tests for selection, so to locate the statistically significant features, the statsmodels library has been used.

```
# (1) select a significance value
alpha = 0.05

# (2) Fit the model
model = sm.OLS(y, X).fit()

# (3) evaluate the coefficients' p-values
model.summary()
```

OLS Regression Results

Dep. Variable:	meantempm	R-squared:	0.945
Model:	OLS	Adj. R-squared:	0.943
Method:	Least Squares	F-statistic:	626.4
Date:	Sun, 04 Apr 2021	Prob (F-statistic):	0.00
Time:	20:18:49	Log-Likelihood:	-1213.1
No. Observations:	676	AIC:	2464.
Df Residuals:	657	BIC:	2550.
Df Model:	18		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-1.2438	0.553	-2.247	0.025	-2.331	-0.157
meantempm_1	0.0882	0.163	0.542	0.588	-0.231	0.408
meantempm_2	0.2695	0.163	1.656	0.098	-0.050	0.589
meantempm_3	0.1839	0.163	1.125	0.261	-0.137	0.505
mintempm_1	0.2278	0.089	2.573	0.010	0.054	0.402
mintempm_2	-0.1240	0.090	-1.372	0.170	-0.301	0.053

Fig. 7. OLS Method

We will find the highest p-value predictor and compare if it is greater than our selected alpha. After finding the predictor we will drop it from our data frame. We will start by bringing in the train_test_split() work from sklearn.model_selection module and using SciKit-Learn to divide our dataset into testing and preparation sets. To ensure you receive a similarly erratic determination of information, we will divide the preparation and testing datasets into 80% preparation and 20% testing and use a random state of 12. This random_state barrier is really useful for ensuring that findings can be replicated. We determined that the regressor model can account for around 95% of the change seen in the outcome variable, mean temperature, using its score() capability. Additionally, we used the sklearn. metrics module's mean_absolute_error() and median_absolute_error() to establish that, on average, the predicted value is wrong by around 1.11 degrees Celsius, and that, half of the time, it is off by about 0.92 degrees Celsius.

B. DNN Regressor Model

After applying linear regression we will build a DNN Regressor model for our dataset. We have removed the maxtempm and mintempm columns because they are useless in predicting the average meantemp, and predicting the future so we obviously can't have information about what's to come. We have separated the targets(Y) and features(X). The datasets are divided into testing sets and training sets. So to display a better iterative process of training this neural network we are using the validation set as an additional dataset. The training set has 80 % of the data while the testing set and validation set has 10 % of the remaining data. For splitting of dataset scikit learns module to train_test_split() is used.

We have defined a function i.e. reusable which we can refer to as an input function that will call wx_input_fn( ). For feeding the data into the neural network during the testing and training phases we will use this function. Once we have defined our input function then we can train our model on the training dataset. It gives the output for the total loss and average loss i.e. the Sum of Square Errors and Mean Squared Error respectively for the training step which for this one is the 260th step.

There are collection of evaluations for each iteration and we plot them as training steps of function so that we can validate that we have not overtrained our model therefore we will use the matplotlib pyplot module to plot a simple scatter plot. We confirmed that the model can explain around 86% of the variation seen in the outcome variable, mean temperature, using the regressor model's score() capability. In addition, we used the mean_absolute_error() and median_absolute_error() functions of the sklearn. metrics module to verify that, on average, the predicted value is 1.68 degrees Celsius off, and that, around half the time, it is off by 1.10 degrees Celsius.

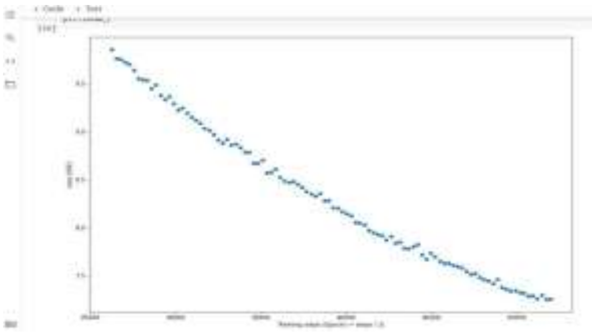

Fig. 8. Graph of evaluation

IX. Conclusion & future scope

Weather prediction is indeed an exigent task since it is extremely dynamic, irregular, and prone to undulations and oscillations which often results in lower prediction accuracy and hence less reliability. To avoid and mitigate these credibility issues one must consider an exemplary dataset along with a suitable corresponding modal as per the need of the hour.

The concerned paper aims to predict the rainfall estimations conditions using historical data with a focus on improving the accuracy of prediction using machine learning techniques. Based upon the acknowledgment extracted from all the research papers, We look at long short-term memory (LSTM) and artificial neural networks (ANN) temporal modelling methods and provide a novel, lightweight, data-driven weather forecasting model. We compare its performance with existing classical machine learning methodologies.

We have made a big discovery as a result of our research phase: machine learning technology may provide intelligent weather forecasting models that are more efficient and promising than currently used conventional approaches. This insight highlights the possibility of machine learning to offer straightforward and upbeat solutions for precise weather predictions. The research only gets us to the conclusion that it is more effective to use the deep volume of data to its maximum potential in the territory of prediction. The integration of machine learning methods with accessible Internet of Things (IoT) gadgets, such temperature and humidity sensors, may become a key topic of attention in the future. The collection of meteorological information from numerous sites inside a city is made possible by this combination, creating a substantial vertical for future growth.

References

[1] E. N. Lorenz, "Empirical orthogonal functions and statistical weather prediction," vol. 1, Cambridge: Massachusetts Institute of Technology, Department of Meteorology, 1956.

[2] Y.-K. Wu and J.-S. Hong, "A literature review of wind forecasting technology in the world," in 2007 IEEE Lausanne Power Tech, 2007, pp. 504-509.

[3] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," Philosophical Transactions of the Royal Society A, vol. 379, no. 2194, p. 20200209, 2021.

[4] G. Hoogenboom, "Contribution of agrometeorology to the simulation of crop production and its applications," Agricultural and forest meteorology, vol. 103, no. 1-2, pp. 137-157, 2000.

[5] M. Kassab et al., "Exploring research in blockchain for healthcare and a roadmap for the future," IEEE Transactions on Emerging Topics in Computing, vol. 9, no. 4, pp. 1835-1852, 2019.

[6] A.G. Salman, B. Kanigoro, and Y. Heryadi, "Weather forecasting using deep learning techniques," in 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2015, pp. 281-285.

[7] S. Pandey et al., "Weather Forecast through Data Mining," International Journal of Scientific Research in Computer Science, Engineering and Information and Technology, vol. 7, no. 3, pp. 90-95, 2021.

[8] M. Holmstrom, D. Liu, and C. Vo, "Machine learning applied to weather forecasting," Meteorol. Appl, vol. 10, pp. 1-5, 2016.

[9] R. Samya and R. Rathipriya, "Predictive analysis for weather prediction using data mining with ANN: a study," International Journal of Computational Intelligence and Informatics, vol. 6, no. 2, pp. 150-154, 2016.

[10] E. Abrahamsen, O. M. Brastein, and B. Lie, "Machine learning in python for weather forecast based on freely available weather data," in Proceedings of The 59th Conference on Simulation and Modelling (SIMS 59), 26-28 September 2018, Oslo Metropolitan University, Norway, no. 153, pp. 169-176, Linköping University Electronic Press, 2018.

[11] S. Singh et al., "Weather forecasting using machine learning techniques," in Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE), 2019.

[12] R. Kingsy Grace and B. Suganya, "Machine learning based rainfall prediction," in 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 227-229, IEEE, 2020.

[13] P. Kalaiyarasai and A. Kalaiselvi, "Data Mining Techniques Using Weather Prediction," IJCST, vol. 6, 2018.

[14] P. Hewage et al., "Deep learning-based effective fine-grained weather forecasting model," Pattern Analysis and Applications, vol. 24, no. 1, pp. 343-366, 2021.

[15] B.M. Ehsan et al., "Advanced wind speed prediction using convective weather variables through machine learning application," Applied Computing and Geosciences, vol. 1, p. 100002, 2019.

[16] F. Cassola and M. Burlando, "Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output," Applied energy, vol. 99, pp. 154-166, 2012.

[17] P. Shinde et al., "Study on Crime Examination and Forecasting using Machine Learning," Journal of Pharmaceutical Negative Results, pp. 7586-7594, 2022.

[18] Y. Jia and X.Y. Zhou, "Policy evaluation and temporal-difference learning in continuous time and space: A martingale approach," Journal of Machine Learning Research, vol. 23, no. 154, pp. 1-55, 2022.

[19] R. Huang, X. Tan, and Q. Xu, "Learning to learn variational quantum algorithm," IEEE Transactions on Neural Networks and Learning Systems, 2022.

[20] Z.C. Lipton et al., "Learning to diagnose with LSTM recurrent neural networks," arXiv preprint arXiv:1511.03677, 2015.

[21] D. Barua, "Dhaka Stock Exchange stock price prediction using Machine Learning and Deep Learning Models," Doctoral dissertation, Brac University, 2022.

[22] C. MacLeod, "The synthesis of artificial neural networks using single string evolutionary techniques," PhD diss., 1999.

[23] K.P. Ferentinos, "Neural network fault detection and diagnosis in deep-trough hydroponic systems," Cornell University, 2002.

[24] R.D. Dhaniya and K.M. Umamaheswari, "RETRACTED ARTICLE: Brain tumor identification and classification of MRI images using data augmented support vector machine," Cognitive Neurodynamics, vol. 16, no. 4, pp. 973-973, 2022.

[25] G.M. Foody et al., "Training set size requirements for the classification of a specific class," Remote Sensing of Environment, vol. 104, no. 1, pp. 1-14, 2006.

[26] M. Zakaria et al., "Artificial neural network: a brief overview," neural networks 1, 2014.

[27] N.K. Siavash et al., "Prediction of power generation and rotor angular speed of a small wind turbine equipped to a controllable duct using artificial neural network and multiple linear regression," Environmental research, vol. 196, p. 110434, 2021.