International Journal of Global Research Innovations & Technology (IJGRIT)
ISSN : 2583-8717, Impact Factor: 6.972, Volume 03, No. 02, April-June, 2025, pp 23-28

Data Analytics by Python AI, to Understand Deviation of Minimum Temperature During Winter (Using Artificial Neural Network)

Sumana Chatterjee*

Ph.D. Scholar (Computer Science), Nirwan University, Jaipur, Rajasthan, India.

*Corresponding Author: sumana.spssaha.chatterjee@gmail.com

ABSTRACT

This paper is based on data analysis by Python programming language on Google collaborator platform, supported by the AI code, model used as neural network, to understand the deviation of minimum temperature for the period November to February in comparison with recorded normal. The source of data is online data collection platform 'INDIA METEOROLOGICAL DEPARTMENT, PUNE' and the online site 'OGIMET'. The surface data for Alipore (42807) and the recorded normal temperature, as available there has been collected in csv file format, then uploading the csv file in Google collaborator platform, executed analysis by Python neural network model technique, to understand the predicted output, i.e. pattern of deviation of minimum temperature, based on analysis of big data, with minimum temperature data from 1969 to 2024. In this case the input column used is the column created with data of difference between minimum temperature and recorded normal for that month. The mod value of temperature difference for the four months, November to February filtered from this big data set has been used as filtered and scaled data, subjected to analyse to understand the future trend of this deviation. Model fit has been used with train-test split in 80%-20% ratio. The artificial neural network model actually resembles the structure of human brain, which can perform intelligent tasks similar to human brain by the process of learning by the machine. Like human brain a similar network is there with mapping between input and output, information propagated through different layers, using back propagation technique to learn the model to get best output with minimum loss. This technique resembles to the networking of human brain and neuron, where information is propagated through synaptic joints to react accordingly. In this paper prediction as well as analysis has been made with the application of artificial neural network with used model as LSTM supported by activation function and optimiser to learn the model for best fit output. The loss value of the model has also been verified to understand the success rate of the model.

***Keywords**: Python, AI, Google Collaborator, Train-test, Neural Network Model, Back Propagation Technique, LSTM, Activation Function, Optimiser.*

Introduction

The interest of data analysis to compare the minimum temperature with its normal value during winter season evoked in mind from that very point of nowadays experience that since one or two years, it is noticed that the minimum temperature during winter fluctuates very often causing rise of temperature even during mid winter also. Temperature difference between two consecutive days, some times is noticed to be significant, increased or decreased more than 3-4 degree Celsius and sometimes more than that also. The sudden deviation of temperature during winter season nowadays takes place very often unlike nature of winter days of former years. From this point of observation, a keen interest of performing analytical study has been developed to analyse the weather data set to understand the trend of change of minimum temperature during the months November to February. So for this purpose from the online Data Supply Platform PUNE, collected data set of daily weather summary, data recorded at 0830 hrs IST, different weather parameter data have been obtained such as maximum temperature, minimum temperature, average wind speed, rainfall, sunshine hours, station level pressure, mean sea level pressure, dry bulb temperature, wet bulb temperature, dew point, relative humidity, vapour pressure, wind direction, visibility, cloud and weather phenomena. As for this work, only minimum

temperature data is required, so filtered out only this data for analysis. Thus obtained this data collected for the period 1969 to November 2023, as available in the site of IMD DSP PUNE, till this period. But as for the analytical study, current data till current months of these months for the year 2024 onwards for November, December, January, February are required also, so the same for the year 2024 and January, February for the year 2025 was to collect from another source 'OGIMET', data for almost same recording time.

Literature Review

Some of the research papers, related to weather forecasting by neural network, as available in internet have been studied. Moreover the code as obtained from Google collaborator platform with used model as 'LSTM' supported by activation function and optimiser, has been used with modification as necessary to get the output with necessary algorithm as wanted for probable prediction of temperature difference of minimum temperature during November to February of 2025 along with prediction of trend of temperature difference in mod value. Also the code of RNN (recurrent neural network) obtained from AI and data science course from 'CCE IIT Madras' conducted by 'INTELLIPAAT', has been used to find the trend of variation of minimum temperature during months November to February, departure from normal value to understand future trend. The validation as well as loss have been cross verified also to understand the success rate as well used as second method also apart from LSTM model done by code supported by Python AI. The research papers as studied from net, have been used to get supporting idea of application in this field.

Research Gap

In this case, instead of actual weather parameter, here as trend of minimum temperature under consideration was to be subjected to analysis, for better understanding, how the deviation of minimum temperature will be in near future, the function representing the absolute value of the difference of minimum temperature from the normal value is the variable taken for analysis in neural network. So the factor of analysis is the data of variation instead of the actual weather parameter.

Research Questions /Hypothesis

After obtaining weather data from the online data supply platform DSP-Pune and OGIMET, collected data of minimum temperature and record of normal to compare the absolute value of deviation. Analysis of this deviation by the process of back propagation technique in neural network, firstly with code obtained from generative AI with LSTM model, then code of RNN with LSTM model, helped to figure out the trend of deviation of minimum temperature in future.

Methods

The analysis process has been done with the process of back propagation technique of artificial neural network. Artificial neural network resembles the network of human brain and nervous system. Almost 100 billion interconnected neurons in a human brain are there functioning together to aid decision making parts and functioning. Dendrite takes signals (stimulus) from the other neurons, cell body (soma) processes the signal and may or may not fire the neuron(excitation or inhibition). Axon transmits the outputs to other neurons or cells. Similar as biological neuron, artificial neuron also receives signals through inputs, processes these signals and transmits signals to give output. The whole transmission process is based on mathematical model, mimics the functioning of biological neuron. Takes inputs in the form of numbers, processes this to get output. Different models of artificial neurons have been developed based on this idea. Basically artificial neural network follows the perceptron model. In this model inputs are real numbers, neurons take the weighted combinations of inputs. Bias is added to weighted inputs. Weighted input passes through an activation function to give output. ANN consists of multiple layers with multiple neurons in each layer (hidden layers).Each neuron (except inputs) represents a perceptron model. Every neuron in one layer is connected to every neuron in the successive layer. Output of one neuron are passed as inputs to the next layer. Similar as human brain, only one neuron can not process complex tasks. Multiple neurons in different layers can process the whole tasks together, supported by activation function. Activation function is like a gate between the input and output of a neuron, to introduce non-linearity into the model and enable learning complex functions (models). Similar as human brain and nervous system, the task of accepting input, transmit this signal to next neuron and getting output, all these performance is not possible by single neuron. The neurons or perceptrons are arranged hierarchically in layers, each of one layer is interconnected to each of the successive layer. Each neuron (except inputs) represent a perceptron model. Output of one neuron are passed as inputs to the neurons of next layer. The weighted inputs along with bias, activated by suitable activation function, propagated through different hidden layers, reached to the output layer to get the output. The whole process again is

repeated several times through the process of back propagation to minimise the loss or error of calculation, by learning of machine, the success of the model depending upon the rate of machine learning, epoch value (number of repetitions) and suitable activation function. The loss value or error is minimised generally with low learning rate, i.e. suitable number of epochs. The low value of learning rate and suitable number of epochs can lead to maximum accuracy of model which can be determined by neural network model and learning of machine by trial and error method, by adjustment of these factors. There are various types of activation function, such as 'Linear activation function', 'Sigmoid activation function', 'Tanh' activation function, 'Relu' activation function, 'Softmax' activation function. The application of suitable activation function depends on the choice of field for analysis and nature of parameters. In this paper, Recurrent Neural Network simulation process has been applied to get the output with minimised loss. RNN is called as recurrent because they perform the same task for every element of a sequence, with the output being dependent on the previous computations. Another way to think about RNN is that they have a memory which captures information about what has been calculated so far. Back propagation through time is used to update the weights in the recurrent neural network. RNN typically predicts one output each time step. Conceptually, back propagation through time works by unrolling the network to get each of these individual time steps. Then, it calculates the error across each time step and adds up all of the individual errors to get the final accumulated error. Following which the network is rolled back up and the weights are updated. The equation of long term dependency can be defined as $\partial E/\partial W = \partial E/\partial 3 * \partial 3/\partial h3 * \partial h3/\partial 2 * \partial 2/\partial h1$…. There arises a long dependency while back propagating the error. Now, if there is a really long dependency, there's a good probability that one of the gradients might approach zero and this would lead to all the gradients rushing to zero exponentially fast due to multiplication. $\partial E/\partial W=0$, Such states would no longer help the network to learn anything. This is known as vanishing gradient problem. LSTM or long short term network is special kind of RNN which is explicitly designed to avoid the long term dependency. The key to LSTM is the cell state. The cell state is kind of like conveyor belt. It runs straight down the entire chain with only some minor linear interactions. It is very easy for information to just flow along it unchanged. The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates. Gates are a way to optionally let information through. They are composed of a sigmoid net layer and a pointwise multiplication operation. The sigmoid layer outputs numbers between zero and one describing how much of each component should be let through. A value of zero means let nothing through and value of one means let everything through. The first step in LSTM is to decide what information we're going to throw away from the cell state. This decision is made by a sigmoid layer called the "forget gate layer". The next step is to decide what new information we're going to store in the cell state. This has two parts. First, a sigmoid layer called the "input gate layer" decides which values we'll update. Next, a tanh layer creates a vector of new candidate values, that could be added to the state. Then we have to update the old cell state, into new cell state. In this paper LSTM model (one kind of recurrent neural network as described before) has been applied here to determine the trend of departure from normal for the minimum temperature. The process of analysis has been done by generative AI in google collaborator platform as well as cross check by RNN python code with the use of same LSTM model. Initially code supported by AI, google collaborator, code started with importing packages as usual numpy, pandas, min max scaler, from tensorflow keras model, imported sequential, imported LSTM model for analysis and also imported matplotlib pyplot for plotting chart. Then uploaded the file for analysis already having in content folder of google collab, in this case the file with data with minimum temperature of months November to February for years 1969 to 2024. The data for January 2025 and February 2025 have been obtained in the data set also. The data till November 2023 is available in data collection platform IMD Pune, the data after this period, till February 2025, is collected from OGIMET, online meteorological data collection platform. The normal temperature for the months subjected to analysis, the recoded normal for the months of November, December, January, February are mapped accordingly to compare with the actual minimum value. The mod (absolute) value of the difference between actual minimum temperature and recorded normal of that day of that month, is taken as the parameter of analysis. Completed the necessary feature engineering, such as converting the minimum temperature field to numeric field, drop null values from the data set. Then scaled the column of 'mod value of temperature difference' by 'Min Max scaler()'. After this created sequential model LSTM using scaled data of the column with mod value of temperature difference. Then created the training and testing set, in the ratio 80%20%. Then built the LSTM model with activation 'tanh'. Instead of 'tanh', the model had also been executed by 'relu' activation function and noticed a little change only. Lastly compilation of the model was done with used optimiser as 'adam' and loss as 'mse'(mean square error). Then fitted the model with suitable epochs and predicted model output with test data to get minimum loss accordingly. Then with the input of this big data set, predicted

the probable trend of temperature difference from normal for the year 2025.To cross check the model and also to get model output by another neural network, here performed recurrent neural network (RNN) by other code, not provided by generative AI, with the same data set and plotted the loss function, validation and test data to understand the success of model. The equation of neural network is defined by the function f, defined as $f(b+^{n}\sum x_i w_i)$, where b is the bias, x is input to neuron, w is weight. Number of inputs is n, i counter, 1 to $n.^{i=1l}$

Significance of the Study

Net special kind of RNNs which are explicitly designed to a from this data analysis, on the basis of analysis of big data with meteorological weather parameters, in this case minimum temperature recorded for daily summary data of Alipore (42807), obtained from the data collection platform, IMD Pune and also minimum temperature recorded, from daily summary data of OGIMET, all these minimum temperature data for four months November to February, were compared with recorded normal value, obtained from years normal, available in data collection platform IMD PUNE. From this comparison, the difference, mod(absolute)value was taken under consideration for analysis to understand the trend of departure from normal. From this type of analysis the trend of change of normal weather pattern, can be determined for understanding the weather situation in near future. In this case, difference of minimum temperature from recorded normal value for four months November to February for the years 1969 to 2024 and January, February for 2025, was considered for analysis by neural network model, by code obtained from GEN AI, as well as by python code in google collaborator platform. For understanding winter temperature data November to February, this span of year has been taken under consideration.

Timeline

Time taken to collect the data from website of data collection platform and OGIMET, time to make compatible with adequate data columns and also to make the python code ready for data analysis in google collaborator platform. Time taken to check compatibility of parameters, validity of model and success rate.

Conclusion and Future Work

Similar type of analytical study with other meteorological parameters to understand deviation from normal may also be done in future to understand pattern of change of weather and climate. Different types of research studies are there nowadays to understand the changing pattern of environment, weather pattern, etc., in this way similarly this study was a try also by neural network model with non-linear data to understand the pattern of change of minimum temperature of four months in year, November to February, as this could be taken as the period of study of minimum temperature of winter season. Other types of study with other weather parameters, such as rainfall, maximum temperature etc. can also be made similarly.

Image of Neural Network and Some Screenshots of Software, as Used

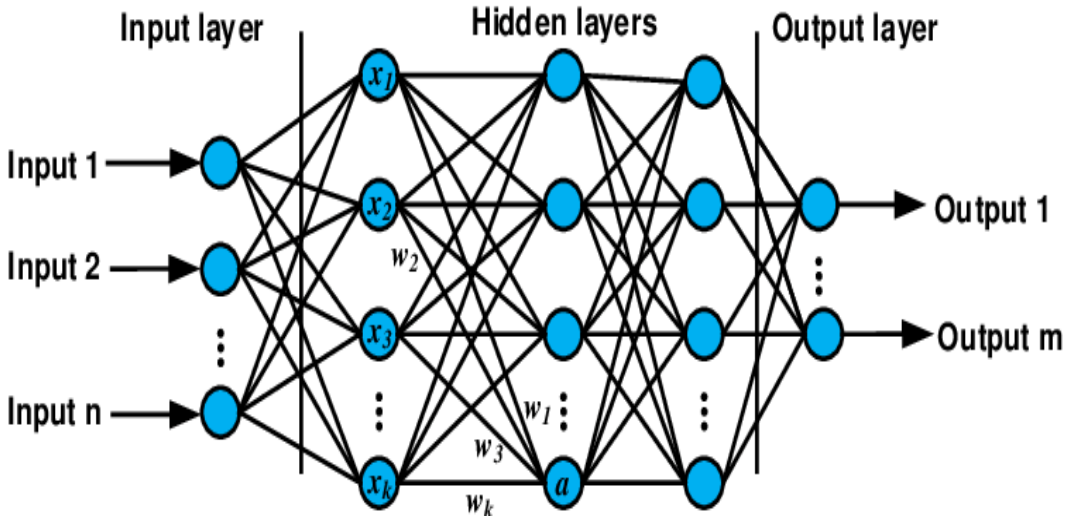

Image 1: Image of Neural Network

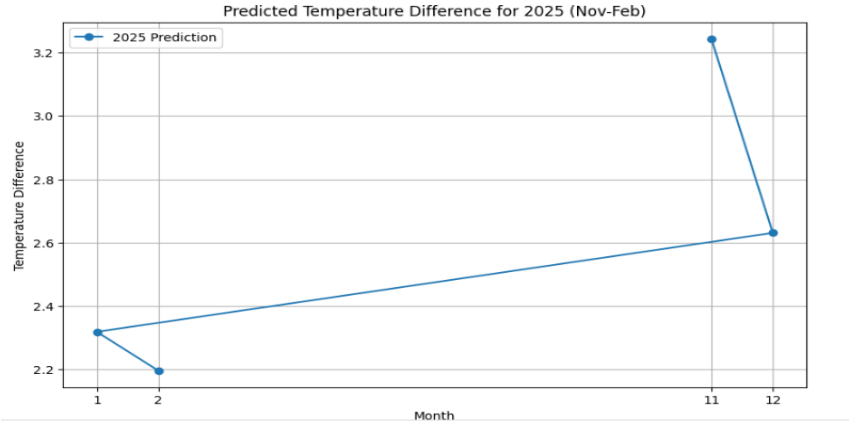

Image 2: Image of Probable Temperature Difference (2025)

```
plt.subplot(2,1,1)
plt.title("Loss & Val Loss")
sns.lineplot(history_loss,palette="flare");
```

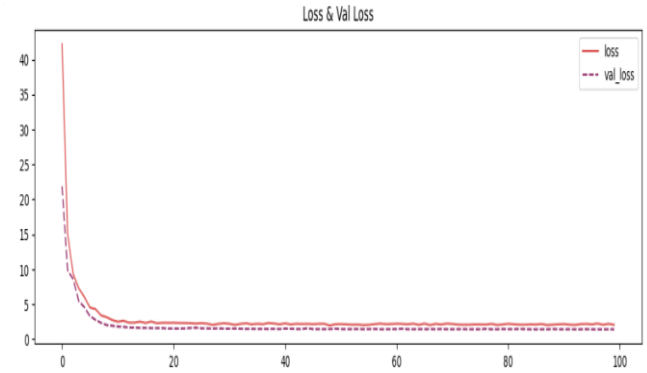

Image 3: Image of Loss (MSE)

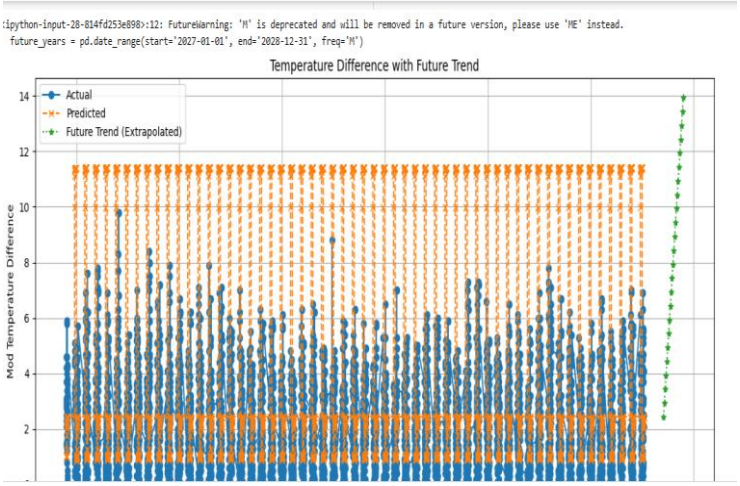

Image 4: Image of Temperature Difference with Future Trend

References

1. Baboo, S. S. et al.(2010, october). An Efficient Weather Forecasting System using Artificial Neural Network. *International Journal of Environmental Science and Development, Volume 1* (4), IJESD 2010 Vol.1(4): 321-326 ISSN: 2010-0264. doi:DOI: 10.7763/IJESD.2010.V1.63
2. Kakar, S. A. et al.(2018). Artificial Neural Network based Weather Prediction. *(IJACSA) International Journal of Advanced Computer Science and Applications, 9*(8). doi:DOI: 10.14569/IJACSA.2018.090859
3. Olaiya, F. et al.(2012). Application of Data Mining Techniques in Weather Prediction and Climate Change Studies. *I.J. Information Engineering and Electronic Business, 2012, 1, 51-59*, 51-59. doi:DOI: 10.5815/ijieeb.2012.01.07
4. Tyagi, H. et al.(2016, April). Weather - Temperature Pattern Prediction and Anomaly Identification using Artificial Neural Network. *International Journal of Computer Applications (0975 – 8887), Volume 140* (3), 15-21.
5. Narvekar1, M. et al.(2015, July). Daily Weather Forecasting using Artificial. *International Journal of Computer Applications (0975 – 8887), Volume 121* (22), 9-13.