IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Detailed Survey Of Weather Prediction Using Machine Learning Algorithms

Dr. H S Prasantha
Professor
Department of Computer Science and Engineering
K.S. Institute of Technology, Bangalore, India

Abstract: Perfect weather predictions are needed for a plethora of activities and it was one of the main challenging problems the entire world faced because of its multiple dimensions and non-linear trends. Weather depends on multiple climatic conditions like temperature, air pressure, humidity, wind flow speed and direction, cloud height and density, and rainfall. The most common phrases/keywords in weather prediction related articles were 'wind', 'precipitation', 'climate change', 'wind forecasting' and 'ensemble prediction'. The most common countries in which surveys were done are China, USA, Australia, India and Germany. Extreme meteorological events are often related to the occurrence of weather fronts.

Keywords- Weather, machine learning, prediction

1. Introduction

Weather forecasting is the application of scientific techniques and technology to predict the conditions of the atmosphere at a certain location and time. For a long time, the researcher had attempted to establish a linear relationship between the input weather data attributes and the corresponding target attribute. But the discovery of nonlinearity within different attributes of weather data, the focus has shifted towards the nonlinear prediction of the weather. Weather forecasts are made by collecting quantitative data about the current state and previous trend of the atmosphere and using scientific understanding of atmospheric processes to predict how the atmosphere will evolve. The weather warning is important for the protection of life and property. Weather prediction has seen a variety of approaches in recent years based on Genetic Algorithms and Neural networks but these fail to capture the complex relationships between various factors which affect weather.

Machine Learning is a part of Artificial Intelligence with the help of any system can learn and improve from existing real datasets to generate an accurate output. The main techniques that are used for weather prediction is unsupervised learning, supervised learning, Support Vector Machine, FP Growth Algorithm and Decision Tree Classification. Machine learning in weather forecasting is a recent trend in the literature. There are several works which discuss this topic. Holmstrom et al. proposed a technique to forecast the maximum and minimum temperature of the next seven days, given the data of past two days. They utilized a linear regression model, as well as a variation of a functional linear regression model. They showed that both the models were outperformed by professional weather forecasting services for the prediction of up to seven days. However, their model performs better in forecasting later days or longer time scales. A hybrid model that used neural networks to model the physics behind weather forecasting was proposed. Support vector machines were also utilized for weather prediction as a classification problem.

There are three important aspects to be taken into account when planning work using machine learning methods in Numerical Weather Prediction (NWP) models. The first is to speed up computations of very computationally expensive parts of the model, the second is to improve the performance of current algorithms, and the third is to emulate the existing code with machine learning models in order to easily allow a model to run on a computer cluster with GPU accelerators. Currently used databases are usually available with very coarse resolution and consist of numerous errors. Convolutional Neural Networks (CNN) can be used to improve them with the use of Sentinel-2 satellite data, the CORINE land-cover, and the BigEarthNet database.

The transfer functions normally used in multilayer neural network are tansig, logsig. And purelin. Tansig is a neural transfer function used to calculate a layer's output from its net input. It is used because meteorological observed data is nonlinear in nature, from which it is difficult to generate the frequent patterns of similar type of data related to one of the output classes. This activation function gives output from -1 to +1. Purelin is a Neural Transfer Function used in fitting problems. This activation function is used in the middle layer between two hidden layers.

There are two types of Neural Networks: Static and dynamic. Multilayer Perceptron (MLP) and Radial Basis Function (RBF) are types of Static Neural Networks while Time Delay Neural Networks (TDNN), concentration-in-time neural networks are dynamic neural networks. All these models do a good job in identifying the seasonal variations but fail in trend and random variations. A variety of Time Series Models have also been developed like Auto regression, Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) etc. but combining time series with the neural network is mostly unexplored. A data mining based predictive model to identify the fluctuating patterns of weather conditions was also proposed. The patterns from historical data are used to approximate the upcoming weather conditions. The proposed data model uses Hidden Markov Model for prediction and k-means clustering for extracting weather condition observations. Weather prediction was studied via a hybrid approach, which combines discriminatively trained predictive models with Deep Neural Networks that models the joint statistics of a set of weather-related variables. Monitory et al. used the concept of crowd sensing, where participating users share their smart phone data to environmental phenomenon. They introduced an architecture named Sen Square, which handles data from IoT sources and crowd sensing platforms, and display the data unified to subscribers. This data is used in smart city environment monitoring.

2. LITERATURE SURVEY

The efforts are made in the paper to identify the contributions made by different authors related to identification and prediction of Heart related issues. The detailed discussion is made with respect to the approaches/methodology adapted and the outcome of the work.

The paper (1) titled Weather Forecasting Using ANN with Error Backpropagation Algorithm uses ANN with error back propagation. The author Used a dataset of over 7 years. Scope of model was limited to Mumbai. Tansig, logsig and pureline activation functions were used. There were 28 input neurons, 10 in hidden layer 1 and 10 in hidden layer 2 and 5 output neurons. The Error rate= 0.0773 and accuracy = 90% is achieved.

The paper (2) titled Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives uses ‘tidytext’ R package for text mining. The outcome of the work is Text mining and analysis of 500 most relevant articles since 2018 was done. Most common methods in NWP articles were DL, RF, ANN, SVM and XGB. ML is efficient in predicting weather.

The paper (3) titled intelligent weather forecast uses Dynamic weighted time delay neural networks. The contribution is short-term temperature and rainfall forecasting over East coast of China. Applied on 8 different weather forecasting stations. Data was read as an ASCII file. Preprocessing included accounting for missing data, reducing chaotic behaviors, low pass filter and normalization. Error rate for 2 stations =0.7092; 4 stations = 0.7485; 8 stations = 0.8223

The paper (4) titled a novel weather prediction model using a hybrid mechanism based on MLP and VAE with fire-fly optimization algorithm uses A novel weather prediction model using a hybrid mechanism based on MLP and VAE with fire-fly optimization algorithm. The Data was collected from Delhi. VAE was used to process global features and its output was given as input to MLP where local or internal features were extracted. Fire-fly optimization was used for VAE-MLFNN and MLP. Goal was to minimize RMSE. Accuracy for different criteria such as temperature, snow, humidity, pressure, wind speed and visibility varied from 80-95%

The paper (5) titled Daily Weather Forecasting using Artificial Neural Network uses Artificial Neural Networks. The Preprocessing algorithms employed were cleaning, normalizing and splitting into train data and test data. Instances considered were intensity of rainfall and sky conditions. Weight and bias were initialized randomly and adjusted with MSE. Learning rate parameter and sigmoid activation function was used. ANN with back propagation was concluded to be most appropriate for weather predictions.

The paper (6) titled Disease Manifestation Prediction from weather Data Using Extreme Learning Machine uses Proposed ELM (Extreme learning machine) algorithm. The contribution is Predicting when a disease will increase to a threshold that causes significant economic loss is important to prevent. Experiments were conducted for different activation functions and it could give satisfactory accuracy of 91.5% for radial basis function.

The paper (7) titled Survey on weather Forecasting Using Data Mining uses Support Vector Machine (SVM) algorithm. This complete paper narrates different data mining techniques and methodology for weather prediction. Based on result, it can be concluded that the support vector machine algorithm can gives better weather prediction with higher accuracy and provides better result.

The paper (8) titled Real Time Weather Prediction System Using IOT and Machine Learning uses K-nearest neighbor. The temperature, light and humidity are the three important parameters that are monitored and uploaded on thingspeak cloud. A Logistic regression model has been used in Jupyter notebook environment that is trained with prerecorded values of parameters and used to predict the weather parameters in real time environment. The result of the model is also compared with the other works available in literature and the proposed system is slightly better in terms of accuracy. Further, the system can be modified to be used at commercial level and have many applications in smart homes, buildings, sports

The paper (9) titled Weather Forecasting Using Machine Learning Techniques uses Support Vector Machine(SVM), Artificial Neural Network(ANN),Time Series Recurrent Neural Network (RNN). The paper compares the three methods using the parameter root mean square error between actual value and predicted value. The software used is pandas ,NumPy, Keras , Git ,Matplotlib, TensorFlow, Anaconda and google cloud services. Data is collected from different airport weather stations of India. Various attributes like atmospheric pressure, relative humidity, dew point etc are used. It is concluded that RNN gives the least root square error.

The paper (10) titled Smart Weather Forecasting Using Machine Learning: A Case Study in Tennessee uses Random Forest Regression (RFR),Ridge Regression (Ridge), Support Vector (SVR), Multi-layer Perceptron (MLPR), and Extra-Tree Regression (ETR). Real weather data is collected from 10 cities. Root mean square error is measured to evaluate the models. The target variable for this record is assigned as the temperature at the same timestamp of the next day. They compare root mean square error for one city case and ten city case using different regression models. They come to the conclusion that RFR is more accurate in finding the weather as it shows a low RMSE close to 3.0 for ten city case and a high value for one city case.

3. MACHINE LEARNING ALGORITHM DISCUSSION

The machine learning algorithms used in weather predictions thus far are:

- Ensemble Neural Networks
- General Regression Neural Network
- Regression models
- Recurrent Neural Networks
- Multi-Layer Perceptron
- Genetic Algorithms
- Fuzzy clustering
- Artificial Neural Networks
- Back error propagation
- Support Vector Machines
- K-nearest neighbour
- Naïve Bayes Classifier
- Decision tree
- Learning vector propagation
- Deep Learning
- Random forest
- XGBoost
- Principal component analysis

4. TOOLS

- Python programming higher installed in windows 10 or any distribution of Linux.
- Arduino IDE software
- Hardware:MPU6050 sensor
- HC05 Bluetooth module.
- Training of an ANN model using Coding Source in MATLAB.
- Python open source library called TensorFlow1 is used in order to train the neural network.

5. APPLICATIONS

Weather forecasting is very necessary for making strategic plans and executing daily tasks in multiple application areas like

- Airport: As Airport traffic is vastly reliant on weather conditions, weather forecasting is very necessary for efficient air traffic control.
- Agriculture (Crop growth and production): Crops grown anywhere are majorly dependent on the weather and therefore help in profitable and healthy growth.
- Electricity: Electricity is produced in various ways through natural forces such as tidal, wind and water, monitoring them closely using the weather forecast can help in efficient use of the grid.
- Water reservoir: The volume of water stored can be regulated according to the rainfall prediction.

- Tourism: People's preference of climate is very subjective, with accurate weather forecasting they can choose the time to visit places in their desired weather.
- And daily life

Applications also include

- Climate monitoring: Climate refers to weather conditions over a large area, the average weather conditions help in monitoring the climate.
- Drought detection: Farmers can accurately determine the period over which the crops can be grown.
- Severe weather and flash flood warnings: These natural conditions can be very fatal, timely warnings can be put out to avoid catastrophes.
- River monitoring and control: Rivers can have multiple subsidiaries and canals can also be constructed, water flow can be rerouted and regulated so that floods and droughts can be avoided.
- Numerical weather prediction, model initialization and verification.

Governments and businesses are utilizing their predictions to safeguard life and property and boost operation quality and schedule various day-to-day operations for individuals. Farms, airlines, hospitals, fire services etc. can provide an hour's or even a day's notice with accurate weather forecasting

6. REFERENCES

1. Siddharth Singh, Mayank Kaushik, Ambuj Gupta, Anil Kumar Malviya, "Weather Forecasting Using Machine Learning Techniques",CSED,KNIT
2. A H M Jakaria, MdMosharaf Hossain, and Mohammad Ashiqur Rahman. 2018. Smart Weather Forecasting Using Machine Learning: A Case Study in Tennessee. In Proceedings of ACM Mid-Southeast conference (Mid-Southeast'18). ACM, New York, NY, USA, 4 pages.
3. Bochenek, B.; Ustrnul, Z. "Machine Learning in Weather Prediction and Climate Analyses—
4. Applications and Perspectives" Atmosphere 2022, 13, 180. https://doi.org/10.3390/atmos13020180
5. LL. LAI, H. BRAUN, Q.P. ZHANG, Q. WU, Y.N. MA, W.C. SUN, L.YANG, "INTELLIGENT WEATHER FORECAST", Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004
6. Vuyyuru, Veera& Rao, G. &Srinivasa Murthy, Y.V.. (2021). A novel weather prediction model using a hybrid mechanism based on MLP and VAE with fire-fly optimization algorithm. Evolutionary Intelligence. 14. 10.1007/s12065-021-00589-8.
7. MeeraNarvekar, PriyancaFargose and DebajyotiMukhopadhyay, ""Weather Forecasting Using ANN with Error Backpropagation Algorithm", Proceedings of the International Conference on Data Engineering and Communication Technology, 2017, Volume 468
8. MeeraNarvekar, PriyancaFargose, "Daily Weather Forecasting using Artificial Neural Network", International Journal of Computer Applications (0975 – 8887), Volume 121 – No.22, July 2015
9. R.K. Arora, S. Sharma and B.P. Singh, "Late blight disease of potato and its management", Potato Journal, vol. 41, no. 1, 2014.
10. E. Cambria, G.B. Huang, L.L.C. Kasun, H. Zhou, C.M. Vong, J. Lin, et al., "Extreme learning machines [trends & controversies]", IEEE Intelligent Systems, vol. 28, no. 6, pp. 30-59, 2013.
11. S. Ding, X. Xu and R. Nie, "Extreme learning machine and its applications", Neural Computing and Applications, vol. 25, no. 3-4, pp. 549-556, 2014.
12. Dr. H S Prasantha, "NOVEL APPROACH FOR IMAGE COMPRESSION USING MODIFIED SVD", International Journal of Creative Research Thoughts (IJCRT), Volume 8, Issue 8, Page 2234-2243, Aug 2020

13. Dr. H S Prasantha, “IMPLEMENTATION OF IMAGE COMPRESSION USING FAST COMPUTATION OF SVD ON DM642”, International Journal of Creative Research Thoughts (IJCRT), Volume 8, Issue 8, Page 2364-2368, Aug 2020
14. Prasantha, H, H Shashidhara, K N B Murthy, and M Venkatesh. "Performance Evaluation of H.264 Decoder on Different Processors." International Journal on Computer Science & Engineering. 1.5 (2010): 1768. Web. 7 Apr. 2013.
15. H. S. Prasantha, H. L. Shashidhara, and K. N. Balasubramanya Murthy. Image compression using SVD. In Proceedings of the International Conference on Computational Intelligence and Multimedia Applications, pages 143–145. IEEE Computer Society, 2007.
16. Gunasheela K S, H S Prasantha, “Compressive sensing for image compression: survey of algorithms”, Proceedings of Emerging Research in Computing, Information, Communication and Applications, ERCICA, Springer publication, Bengaluru, 2018
17. K N Shruthi, B M Shashank, Y. SaiKrishna Saketh, H.S Prasantha and S. Sandya, "Comparison Analysis Of A Biomedical Image For Compression Using Various Transform Coding Techniques", IEEE, pp. 297-303, 2016