# Comparative Study of Machine Learning Models in Weather Forecasting

K V S Krishna Praneeth
Department of Information Science and Engineering
R.V. College of Engineering
Bengaluru

Ishan Jha
Department of Information Science and Engineering
R.V. College of Engineering
Bengaluru

Siddhartha Gupta
Department of Information Science and Engineering
R.V. College of Engineering
Bengaluru

Prof Merin Meleet
Assistant Professor
Department of Information Science and Engineering
R.V. College of Engineering
Bengaluru

***Abstract*—Weather forecasting plays an integral role in our daily lives and influences a wide array of activities. This paper presents a survey on the various statistical approaches to weather prediction. The need for this methodology is governed by the instability and high cost associated with the physical model. A comparative study is performed and the aptness of each model is discussed. Various environmental parameters such as temperature, rainfall, humidity are evaluated with the help of these models. Among the existing machine learning techniques, prominent techniques such as Linear and Functional Regression, Artificial Neural Networks and Support Vector Machines are analysed. The results show that the models compare well against the numerical models and the efficacy of the Support Vector Machines and Neural Networks exceeds that of Regression models.**

***Index Terms*— Functional Regression, Linear Regression, Neural Networks, Numerical Weather prediction, Statistical Learning Theory, Support Vector Machines, Weather Forecasting**

## I. INTRODUCTION

Weather forecasting can be understood as the utilization of the present and prevailing techniques, methods and information to predict the state of atmosphere for a location at any given time. Weather forecasting is the event that follows weather analysis. These two processes together play an integral role in our everyday lives and control our lifestyle. Previously, this forecasting was accomplished by viewing atmosphere as a fluid and applying the concepts of fluid dynamics and thermodynamics. This method produced results that were unreliable as a consequence of external factors affecting the highly sophisticated process. The approach used in the recent times incorporates statistical analysis. This method is robust to external factors as a result of which the latter is significantly reliable as compared to the former. The concept of machine learning is the alternative to the existing models. The analysis and the prediction or forecasting of weather or the atmospheric state at a later time at a given location can be accomplished by the implementation of several algorithms such as linear regression, Functional regression, Naïve Bayes and Support Vector Machines. The best suitable model is chosen based on the effectiveness of each of the models being considered for the implementation. The analysis for each model is done on the historical weather data. The outcome of this problem statement is to compare alternatives to accomplish the same and hence choose the most cost effective method as the best suitable technique.

## II. TRADITIONAL APPROACH

Fundamentally, there are two approaches to weather prediction. They are Empirical and Dynamical Methods**.** Dynamic Meteorology is the one which is most prevalent in use in the form of the Numerical Weather Prediction Model (NWP) first proposed by the British mathematician Lewis Fry Richardson in 1922. This is the study of atmospheric motions as solutions of the fundamental equations of hydrodynamics of the atmosphere and the mathematical equations which can describe large scale movements in the atmosphere. In this paper [9], the theory states that the set of equations which defines the state of the atmosphere are governed by the laws of fluid dynamics and thermodynamics so as to predict how the fluid behaves in the future. It is a robust mathematical model which takes the present weather conditions and processes it to build a model for predicting the weather.

The data is initially gathered by remote sensing satellites. Components of temperature, humidity, wind speed and atmospheric density are among the parameters of weather which are projected using images and data taken by these meteorological satellites and this is interpolated to a geographical grid. These models which approximate the state of the atmosphere are then used to describe how these parameters change over time. Weather parameters and the partial differential equations explaining the current state of the atmosphere are provided to the computer as inputs.

In this paper [8], an initial state of atmosphere is defined and the change in these atmospheric variables is calculated. This process is repeated iteratively and the input to the current phase is provided from the output from the previous phase and this is repeated until the stage where the weather needs to be forecasted. Parameterization needs to be performed to account for the minor atmospheric processes not represented by any equation. A wide assortment of models explaining the behavior of the atmosphere exist. Each model serves a specific purpose and understands the atmosphere in a unique manner. An aggregation of maps, charts and outlines are created with the help of the algorithms and subject knowledge of the meteorologist. These are used in the formulation of surface and upper air maps. A prognostic chart is hence created which is then used in weather forecasting.

However this model comes with its share of limitations. The partial differential equations which are used to calculate the state of the atmosphere at a particular instant are highly turbulent and are dependent on an extensive array of atmospheric micro processes which leads to an incomplete understanding of the model and limited forecasting ability. Due to the large quantum of data generated and the complex calculations involved requires powerful supercomputers.

The Empirical model on the contrary uses the relationship between various weather variables and historical data to predict the weather. Examples include Linear Regression model, Fuzzy Logic, Support Vector Machines and Artificial Neural Networks. This paper [10] delves into the various empirical models and compares the efficacy with the traditional dynamic model.

## III .LINEAR REGRESSION AND FUNCTIONAL REGRESSION

Holmstrom et al. [2] used linear regression and functional regression algorithms to predict the weather. The features used were maximum temperature, minimum temperature, mean atmospheric pressure, mean humidity and weather classification for each day during the period 2011-2015 for Stanford obtained from weather underground.

TABLE I. Sample data with serial number, feature and sample values.

| Sl. No | Feature | Value |
|---|---|---|
| 1 | Classification | Sunny |
| 2 | Maximum Temperature | 96 |
| 3 | Minimum Temperature | 60 |
| 4 | Mean Humidity | 52 |
| 5 | Mena Atmospheric Pressure | 30.29 |

The weather classification data of each day was not used by linear regression as the algorithm cannot be used for classification data. Hence, the inputs comprised of eight features for 2 consecutive days, excluding the classification feature. The algorithm was designed to accurately forecast the highest and lowest temperatures for the following week. In total, 14 features were predicted for each pair of consecutive days. The prediction was made using the formula:

$$\mathrm{H}\theta(x^{(i)}) = \theta^{T}\mathrm{x} = y^{(i)}$$

where $h\theta(x^{(i)})$ is the prediction of $y^{(i)}$ for any $x^{(i)}$ where $x^{(i)}$ is the 8-Dimensional input feature vector, $y^{(i)}$ is the 14-Dimensional output feature vector and θ is the parameter corresponding to x for the $i^{th}$ pair of consecutive days.

A variation of functional regression was the second algorithm used by Holstrom et al. [2] which explored the historical weather data to find patterns similar to the current weather patterns and in turn making forecasts built from the historical patterns. It used feature vectors for the first two days as inputs. The outputs were again maximum and minimum temperature for any of the days following the first 2 days. To each feature, a weight was assigned to normalize the weightage given to each of the features. The root mean square error was used to calculate the standard deviation of the individual error terms. The results showed that linear regression out performed functional regression. The rms error for linear regression was lesser than that of functional regression. The professional weather forecasting service out performed both linear as well as functional regression techniques and the discrepancies reducing over a longer period of time.

T.R.V.Anandharajan et al. [3] also used linear regression algorithm to predict weather. The inputs were weather data and labelling the data to different climate. The methodology used was similar to Holstrom et al. [2] using hypothesis function, cost function and a gradient descent and normal equation additionally. The final intended output was maximum temperature, minimum temperature, rainfall and the weather classification which involved multiclass classification and hence the use of logistic regression. The visualization of results was done using MATLAB. The results were obtained with 90% accuracy.

S. Prabakaran et al. [11] proposed a modified model of linear regression to predict rainfall in various districts of South India. The model entailed applying linear regression to the training data with rainfall as the dependent variable and the cloud cover and average temperature as the independent variables. The error percentage is calculated upon comparing the predicted output value with the actual output value. A part of the error percentage is then added to the training data and multiple iterations are performed until a satisfactory output value is obtained and this is applied to the test data. An average error of 7% is obtained upon implementation of this model.

Sanyam Gupta et al. [12] compared the optimization techniques of Normal Equation method and Gradient Descent method which they used in accompaniment with the linear regression model in order to predict three weather parameters viz. Humidity, Temperature and Dew Point. The obtained results demonstrated that the normal equation method forecasts the weather with a low mean square error value whereas the Gradient Descent technique gave highly errant values for all three parameters.

## IV.ARTIFICIAL NEURAL NETWORK

Artificial Neural Networks are software implementations which resemble the human nervous system. The ANNs comprises of inputs (synapses) which when multiplied by weights (strength of signals) and computation by a mathematical function which governs the activation of a neuron where multiple artificial neurons work together to produce an output with minimal error.

Kumar Abhishek et al. [5] used Artificial Neural networks to predict weather. The tool used to perform the analysis was Neural Network Fitting Tool GUI called nntool which is a part of MATLAB. The goal was to predict individual features such as maximum temperature, minimum temperature, rainfall etc. The input data comprised of 365 samples of data from the past 10 years. The model consisted of 3types of layers which are an input layer followed by a hidden layer and an output layer. Levenberg-Marquardt algorithm was used to train the model. They used an iterative process in which corrective adjustments were made in each step until a desired value was obtained. The trained model was tested to deduce the performance of the model.

Observations were made with different transfer functions, varying number of hidden layers and distribution of neurons among them. From the experiments performed, it was observed that the crucial parameters involved were learning rate of the model, number of neurons per hidden layer, transfer function for hidden layers, number of samples, number of hidden layers and overfitting. It was concluded that an artificial neural network is a better alternative to its counterparts such as linear and functional regression as non-linear dependencies can be expressed in the former but not in the latter.

Meera Narvekar et al. [6] compared various techniques in neural networks such as Artificial Neural Networks, Backpropagation networks, fuzzy neural networks etc. Daily weather has multiple parameters that are not linear, but they need to be processed together to deduce features such as temperature, rainfall, humidity and hence Artificial Neural Networks could be used as they are associated with non-linear data. The backpropagation algorithm was used in parallel with the ANN to minimize the errors produced by the network. The accuracy of the model was deduced by finding the mean squared error between the original result and the corresponding forecast result and thereby extract performance. It was concluded that Artificial Neural Networks with backpropagation was the most appropriate technique for weather forecasting.

Mohsen Hayati et al. [7] used the Neural Network architecture of Multilayer Perceptron to design short-term temperature forecasting systems. The model was trained and tested with 10 years of data. The Sigmoid Transfer function and pure linear transfer function were used for the hidden layers and output layer respectively. Scaled Conjugate Descent (SCG) algorithm was used to train the model instead of backpropagation and Levenberg-Marquardt algorithm due to its faster learning rate and also due to better results produced by the SCG algorithm.

.

The results obtained denoted a reasonable performance and prediction exactness of the model. It was concluded that Multilayer Perceptron could be an important tool for the design of short-term temperature forecasting systems.

## V. SUPPORT VECTOR MACHINES

An alternate approach at weather forecasting using the statistic learning theory is done with the aid of Support Vector Machines (SVM). A group of mathematical functions are used by the SVM algorithms which are collectively called the Kernel. A low dimensional input space is transmuted to a higher dimension input space with the help of a linear separating hyperplane. It can be used both for problems of classification and regression. The Estimation function of the SVR is defined as follows:

$$f(x) = (w(\emptyset\ (x)) + b$$

where $w$ is the weight and $b$ is the threshold which is calculated from the data, and $(\emptyset\ (x))$ is the Kernel function. The aim is to map f($x$), the nonlinear function to a higher dimension feature space.

Radhika*et al*. [4] compared the performance of SVR vs. a Multilayer Perceptron (MLP) for predicting atmospheric temperature. The preprocessing was done by substituting missing temperature values with that of the mean temperature for that month. The optimal length of the span $n$ was derived by experimentation and the maximum temperature of previous $n$ days was used in the prediction of Maximum Temperature for a day. With the measure of the performance as mean square error (MSE), they showed that the SVM which used a Radial Based Kernel Function outperformed the MLP trained by the backpropagation algorithm for all orders and the SVM was more sensitive to the selection of the right parameters.

Perez – Vega *et al.* [15] studied the use of SVM to forecast temperature. The input data consisted of historical records that provided information about environmental factors such as temperature, humidity, clouds, solar radiation, wind speed and precipitation etc. Information related to temperature and wind speed were selected as parameters during pre-processing and a target value for the test data was calculated with the help of different Kernel functions with mean square error (MSE) as a standard of performance. A comparative study was performed upon which they concluded that the Polynomial Kernel had the least MSE followed by the Linear Kernel and RBF Kernel.

The methodologies used by Trafalis et al. [14] were compared with the traditional regression method and existing meteorological equations for rain rate to estimate rainfall using WSR-88D radar data which contains information about spectrum width, reflectivity and velocity wherein only reflectivity was used as a parameter. Estimation of the rainfall rate was solved with the LS-SVR which used a polynomial kernel and this was excelled by the linear regression model and the SVR which used a Gaussian Kernel with Mean Square Error (MSE) as the performance measure. The estimation of rainfall was performed using both the SVR and the LS-SVR and it was observed that the SVR had a higher degree of accuracy when compared to the LS-SVR. The outcomes demonstrated that the techniques of LS-SVR and SVR are superior to traditional regression and rain rate formula used in meteorology. Furthermore, they also noted that the accuracy of the SVR was as good as the best Neural Network architectures tested but required parameter optimisation and selection of the right Kernel function. [13]

## VI. DISCUSSIONS

Results of Holmstrom et al [2] showed that linear regression out performed functional regression. The rms error for linear regression was lesser than that of functional regression. This could be due to the fact that temperature could be predicted based on the data of 2 days but it was not enough to predict the trend in weather pattern as a whole. The professional weather forecasting services out performed both linear as well as functional regression and the discrepancies reduced over longer period of time. This was because the present conditions can be accurately measured by the professional services but not by the algorithms but for a prediction of more than a few days ahead, the discrepancies reduced as the machine learning techniques are robust to disturbances whereas the physical models used by professional services are not as a result of which errors accumulate quickly over longer periods of time. Anandharajan et al. [3] performed the same with linear regression using similar methodology as Holmstrom et al. [2] along with the use of gradient descent. The accuracy obtained was over 90%.

The Regression technique with optimisation techniques for weather estimation by Gupta et. al was not as efficient in the case of Gradient Descent when compared to Normal equation method and gave a large disparity in error rate. This can be intuitively explained by the fact that the Gradient descent requires a large number of iterations to derive the learning rate and hence the best fit and is computationally expensive.
The work done by Prabakaran et al meanwhile showed that a modified form of regression gives a more accurate value when compared to the results obtained by just simple linear regression, this is done by iteratively adding a portion of the error percentage to the training set so that the model gets trained accordingly to give more accurate values and a small portion of 20% of the error rate is chosen intuitively so that the problem of overfitting does not occur and generalisation of the model is upheld.

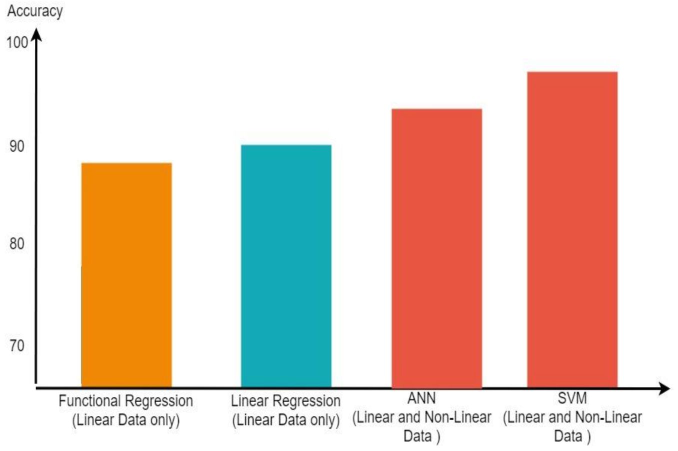


FIGURE I. Comparison of accuracy of models.

Although good performance was obtained with regression, regression techniques cannot be applied to non-linear data. Neural Networks is known for its use with non-linear data. The algorithms used to train the model included the Levenberg-Marquardt algorithm, backpropagation algorithm and Scaled conjugate gradient. Levenberg-Marquardt algorithm was used by Kumar Abhishek et al. [5] and it has the fastest learning rate these advantages reduced considerably as the weights and the biases in the network was enhanced. Meera Narvekar et al. [6] used backpropagation algorithm which is an iterative process and hence produces better results as compared to Levenberg-Marquardt algorithm but it is very slow for practical purposes. Hayati et al. [7] used scaled conjugate gradient as the training algorithm which employed step scaling mechanism as a result of which it was faster than numerous second order algorithms. The results obtained using the algorithm was better than the former algorithms.

The three research papers which explicated the procedure involved in using SVR for weather prediction also gave interesting results. Trafalis et al and Shashi et al. compared the SVR with the Multi-Layer Perceptron and had similar conclusions that the SVR gave a better result when compared to Artificial Neural Network based techniques such as MLP due to their innate ability to give a generalised solution for a given problem and due to the challenges faced in determining the global minima to optimise the Neural Network. Apart from this, another important analysis derived was that there is no formal rubric in defining the parameters and Kernel function which is to be used and this is highly dependent on the problem faced. This can be seen in the results of the research conducted by Perez-Vega et al. where they derived the Polynomial Kernel as the most appropriate Kernel in contrast to the RBF Kernel selected by Shashi et al. The performance metric used in all the three researches was the minimisation of the mean squared error (MSE).

## VII. CONCLUSIONS

Weather forecasting has witnessed a sea change in how it has been conducted. A large cost is incurred while utilising a model in Numerical Weather Prediction along with limited forecasting ability which necessitates the advent of new techniques and methodologies. This paper gave a comparative study of the various Machine Learning techniques which were applied on atmospheric and environmental parameters so as to forecast the weather. The accuracy of several models when compared with professional weather forecasting services was discussed. No one model was adjudicated as the best fit for all measurements and its suitability was dependent on the data set as well as the parameters used. There was a general consensus on SVM and ANN outperforming the more simplistic and traditional approach of Linear Regression. The various techniques applying Statistical Learning Theory concepts discussed in our paper could be used in conjunction with the professional weather forecasting services so as to give more accurate results and to have a larger range of future forecast.

## REFERENCES

[1] Nasimul Hasan, Md. Taufeeq Uddin and Karim Chowdhury, " Automated Weather analysis with Machine Learning", Department of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh.

[2] Mark Holmstrom, Dylan Liu, Christopher Vo, "Machine Learning Applied to Weather Forecasting", Stanford University, December 15, 2016.

[3] T R V Anandharajan, G Abhishek Hariharan, K K Vignajeth, R Jijendran and Kushmita, " Weather monitoring using artificial intelligence", International Conference on Computational Intelligence and Networks,2016.

[4] Y.Radhika and M.Shashi, "Atmospheric Temperature Prediction using Support Vector Machines", International Journal of Computer Theory and Engineering, Vol. 1, No. 1, April 2009.

[5] Kumar Abhishek, M.P.Singhm Saswata Ghosh, Abhishek Anand, "Weather forecasting model using Artificial Neural Network", ELSEVIER,Volume 4,2012.

[6] Meera Narvekar, Priyanca Fargose, "Daily Weather Forecasting using Artificial Neural Network", International Journal of Computer Applications (0975 – 8887) Volume 121 – No.22, July 2015.

[7] Mohsen Hayati, and Zahra Mohebi, "Application of Artificial Neural Networks for Temperature Forecasting", International Journal of Electrical and Computer Engineering, Vol: 1, No: 4, 2007.

[8] Folorunsho Olaiya and Adesesan Barnabas Adeyemo, "Application of Data Mining Techniques in Weather Prediction and Climate Change Studies." I.J. Information Engineering and Electronic Business, 2012, 1, 51-59.

[9] Peter Lynch, "The origins of computer weather prediction and climate modelling." Journal of Computational Physics 227 (2008) 3431–3444.

[10] Pinky Saikia Dutta and Hitesh Tahbilder,"Prediction Of Rainfall Using Datamining Technique Over Assam." Indian Journal of Computer Science and Engineering (IJCSE), Guwahati University.

[11] S. Prabakaran, P. Naveen Kumar and P. Sai Mani Tarun," Rainfall Prediction Using Modified Linear Regression.", ARPN Journal of Engineering and Applied Sciences Vol. 12, No. 12, June 2017 ISSN 1819-6608.

[12] Sanyam Gupta, Indumathy K, Govind Singhal, "Weather Prediction Using Normal Equation Method and Linear regression Techniques." International Journal of Computer Science and Information Technologies, Vol. 7 (3), 2016, 1490-1493.

[13] Nicholas I. Sapankevych and Ravi Sankar. "Time Series Prediction Using Support Vector Machines: A Survey." IEEE Computational Intelligence Magazine May 2009.

[14] Theodore B. Trafalis and Michael B. Richman. "Prediction of Rainfall From WSR-88D Radar Using Kernel-Based Methods." International Journal of Smart Engineering System Design 2003.

[15] Abrahán Pérez-Vega, Carlos M. Travieso, José G. Hernández-Travieso and Jesus B. Alonso, " Forecast of Temperature using Support Vector Machines." International Conference on Computing, Communication and Automation (ICCCA2016), 2016.