Weather Prediction using Advanced Machine Learning Techniques

G. Hemalatha¹, K. Srinivasa Rao², D. Arun Kumar³

³ Machine Learning group of Center for Research and Innovation
¹,² Department of ECE
² Department of CSE
¹,²,³ KSRM College of Engineering, Kadapa, Andhra Pradesh.
¹ latha.g@ksrmce.ac.in, ² ksr@ksrmce.ac.in, ³ dasariak@ksrmce.ac.in

Abstract. Prediction of weather condition is important to take efficient decisions. In general, the relationship between the input weather parameters and the output weather condition is non linear and predicting the weather conditions in non linear relationship posses challenging task. The traditional methods of weather prediction sometimes deviate in predicting the weather conditions due to non linear relationship between the input features and output condition. Motivated with this factor, we propose a neural networks based model for weather prediction. The superiority of the proposed model is tested with the weather data collected from Indian metrological Department (IMD). The performance of model is tested with various metrics.

Index Terms—Indian Metrological dataset, Weather Prediction, Neural networks, Pattern Classification.

1. Introduction

Prediction of weather on daily basis plays an important role in taking the efficient decisions [1]. The dynamics in the conditions of weather motivates the researchers to propose efficient models for data prediction [2]. There exist various type of linear weather prediction models [3]. The problem with linear models is that they do not consider the non-linearity in the input data [4]. In the recent years, non linear models were suggested for weather prediction [5]. The basic weather prediction models are classified as data classification, data clustering and data prediction [6]. Weather data classification is termed as predicting the condition of the weather based on the input parameters [7]. The output of prediction model is a class which demonstrates the condition of the weather [8]. In the weather data clustering, the input data is clustered into groups with each cluster having a cluster centroid. The number of cluster centroids are equal to the number of clusters in the data [9], [13]. Weather data clustering is also called as Unsupervised learning [5], [11], [12]. Weather data prediction is also called as regression analysis in which the value of output variable is predicted using various type of linear and non-linear prediction models [10], [15].

The present study is a weather prediction model related to the weather data classification. In general weather data classification models do not consider the non-linear relationship between the input parameters and the output conditions [13], [14]. Motivated with this factor, a Fully connected neural network (FCNN) model is suggested for weather data classification. The FCNN model considers the non linear nature between the input features and the output class label during the classification stage.
Furthermore, FCNN model has an additional advantage of learning and generalization ability in predicting the class label of the weather data.

2. Proposed Fcnn Model for Weather Prediction

The fully connected neural network is a machine learning model for data prediction. The functional block diagram of FCNN model is given in Fig. 1. The weather sensors of the proposed system acquire the weather data from the environment. The weather data is supervised and labeled before it is divided into training and testing. The training data is fed as input to FCNN model through which the learning process is completed. The test data is fed as an input to the FCNN model to validate the performance. During the validation, the weather condition of the test pattern is predicted.

![Block diagram of FCNN Model](image1)

The detailed architecture of the FCNN model is given in Fig. 2. The architecture of FCNN model consists of input layer, hidden layer and the output layer. The number of input layer nodes is equal to the number of features. The number of hidden layer nodes is greater than or equal to the number input nodes. The number of output nodes is equal to the number of classes in the dataset. The nodes between the layers are connected with the connecting links. Each connecting link is assigned with a weight value. The output at a node is the sum of product of input value and the corresponding weights. The resultant output at a node is passed through the sigmoid activation function. The input values are feed forwarded through the network and the output error at each node of the output layer is calculated. The error at the output layer is back propagated and the corresponding weights are updated using back propagation algorithm (BPA).

![Architecture of FCNN model](image2)

3. Dataset Description

The dataset consists of samples collected on daily basis from the year 2017 to 2020. The feature values of the samples are considered as Temperature, dew point, humidity, wind type, wind speed, wind gust, pressure, precipitation. The output weather conditions were considered as fair, mostly cloudy, partly cloudy, cloudy, light rain and thunder classes. The dataset consists of 1460 samples with 8 features and 6 classes.

4. Results and Discussions

The dataset with 1460 samples is divided into training and testing based on the criteria like 80%, 60%, 40% and 20% for training and the remaining 20%, 40%, 60% and 80% are used for testing.
During the training, the input sample is fed through the network one by one. The architecture of the FCNN model is considered as 8 : 10 : 6 with the number of input nodes representing the number of features, the number of hidden layer nodes greater than the number of features and the number of output layer nodes equal to number of classes in the dataset. The model is trained with the number of iterations (NOIs) = 200. The corresponding mean square error (MSE) is obtained. In general, the MSE will reach to almost nearby zero value. The MSE of the FCNN model for NOI = 200 is given in Fig. 3. The results of FCNN model for various set of training and testing cases is given in Table 1. From Table 1, the performance of proposed FCNN model is highest with 87.83% OA. Similarly, the performance of Fine Gaussian SVM is 63.90% with the less OA. The proposed FCNN model outperformed similar type of models in terms of OA. The histogram for the OA of models is given in Fig. 4. In support to the OA, the performance of proposed model is evaluated with the performance metrics like UA, PA and KC. The advantages of these metrics are that they provide the class wise analysis in classifying the input weather samples. The results of models (a) Fine Gaussian SVM (FGS) (worst case) and (b) FCNN (Best case) with the metrics UA, PA and KC are given in Table 4. In the Table 4, the performance of FGS with UA is 79.63% for class 1 and the performance of FCNN with UA for class 1 is 85.45%. The PA of FCNN model is better than the PA of FGS for class 1. Similarly, the FCNN model outperformed FGS in terms of UA and PA for class 6. The performance of FCNN model is better than FGS with highest KC. The superiority of proposed FCNN model over FGS model is due its learning and generalization ability. In addition, the FCNN model considers the non-linear relationship between the input features of the data and the output class labels.

Table 1: Performance of models (i) Fine Guassian SVM (Worst case) (ii) FCNN (Best case) with the metrics UA, PA and KC for 60% training data.

| Models | Class | UA(%) | PA(%) | KC   |
|--------|-------|-------|-------|------|
| FGS    | Class 1 | 79.63 | 68.24 | 0.867|
|        | Class 6 | 89.14 | 92.46 |      |
| FCNN   | Class 1 | 85.45 | 71.48 | 0.945|
|        | Class 6 | 92.74 | 96.46 |      |

5. Conclusions
In the present study, A FCNN model is suggested for weather data prediction. The proposed FCNN model possesses the learning and generalization ability that captures the non linear characteristics of input features in the dataset. The model outperformed similar type of models in terms of the OA, UA, PA, KC. The model produced the OA of 87.83% as tested with IMD dataset. Furthermore, the model can be extended to the classification of higher dimensional dataset.
Fig. 3. The MSE of the proposed FCNN model for 200 iterations.

Fig. 4. Histogram of the OA for the models considered in the present study.

Acknowledgement
The third author is thankful to the students T. Deepu, U. Siva Sankar, R. Murali and S. Sohail for preparing the dataset.
Table 2: Performance of models for various percentages of training.

| CLASSIFIERNAME                  | TRAINING (%) | OA  |
|--------------------------------|--------------|-----|
|                                 | 20          | 40  | 60  | 80  | 71.38% |
| FineTree                        | 66.34        | 68.23| 70.83| 80.86| 71.38% |
| MediumTree                      | 67.10        | 69.30| 72.70| 81.20| 72.57% |
| CoarseTree                      | 68.10        | 69.50| 72.90| 81.50| 73.00% |
| OptimizableTree                 | 67.50        | 69.90| 74.30| 82.70| 73.60% |
| LinearSVM                       | 65.10        | 68.40| 72.50| 80.10| 73.52% |
| QuadraticSVM                    | 66.40        | 67.90| 72.30| 80.30| 71.72  |
| CubicSVM                        | 65.70        | 68.20| 72.80| 82.20| 72.20  |
| Fine GuassianSVM                | 59.40        | 61.30| 64.50| 70.50| 63.90  |
| MediumGuassianSVM               | 60.10        | 69.20| 74.70| 80.10| 71.20  |
| CoarseGuassianSVM               | 61.80        | 69.30| 73.80| 81.20| 71.65  |
| Fine KNN                        | 58.90        | 60.20| 66.20| 79.90| 66.30  |
| MediumKNN                       | 61.80        | 69.80| 73.80| 85.90| 72.82  |
| CoarseKNN                       | 60.40        | 66.60| 74.20| 86.50| 71.92  |
| CosineKNN                       | 60.10        | 64.30| 72.50| 80.40| 69.30  |
| CubicKNN                        | 62.40        | 65.40| 72.70| 82.60| 70.77  |
| WeightedKNN                     | 60.90        | 62.80| 70.90| 83.80| 69.60  |
| OptimizableKNN                  | 60.60        | 65.90| 74.30| 81.30| 70.50  |
| BoostedTrees                    | 63.40        | 68.70| 72.40| 81.20| 71.40  |
| BaggedTrees                     | 61.20        | 66.70| 74.60| 80.60| 70.10  |
| SubspaceDiscriminant            | 60.20        | 64.30| 73.90| 81.20| 69.90  |
| SubspaceKNN                     | 61.30        | 63.20| 70.60| 82.10| 69.30  |
| RusbootedTrees                  | 53.20        | 57.10| 60.20| 80.30| 62.70  |
| OptimizableEnsemble             | 60.90        | 68.20| 72.90| 83.20| 71.30  |
| FCNN                            | **85.84**    | **87.64**| **88.87**| **89.23**| **87.83**|

References

[1] Abhishek, K., Singh, M.P., Ghosh, S. and Anand, A., 2012. Weather forecasting model using artificial neural network. Procedia Technology, 4, pp.311-318.

[2] D. N. Fente and D. Kumar Singh, "Weather Forecasting Using Artificial Neural Network," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018, pp. 1757-1761, doi: 10.1109/ICICCT.2018.8473167.

[3] Culclasure, Andrew, "Using Neural Networks to Provide Local Weather Forecasts" (2013). Electronic Theses and Dissertations.32.https://digitalcommons.georgiasouthern.edu/etd/32

[4] Litta, A.J., Mary Idicula, S. and Mohanty, U.C., 2013. Artificial neural network model in prediction of meteorological parameters during premonsoon thunderstorms. International Journal of atmospheric sciences, 2013.

[5] Rasp, S., Dueben, P.D., Scher, S., Weyn, J.A., Mouatadid, S. and Thuerey, N., 2020. WeatherBench: A Benchmark Data Set for Data-Driven Weather Forecasting. Journal of Advances in Modeling Earth Systems, 12(11), p.e2020MS002203.

[6] Kumar, A., Liang, P. and Ma, T., 2019. Verified uncertainty calibration. arXiv preprint arXiv:1909.10155.

[7] Föll, R., Haasdonk, B., Hanselmann, M. and Ulmer, H., 2017.

[8] Deep recurrent Gaussian process with variational sparse spectrum approximation. arXiv preprint arXiv:1711.00799.

[9] Trebing, K., Stańczyk, T. and Mehrkanoon, S., 2021. Smaat-unet: Precipitation nowcasting using a small attention-unet architecture. Pattern Recognition Letters, 145, pp.178-186.
[10] Grönquist, P., Yao, C., Ben-Nun, T., Dryden, N., Dueben, P., Li, S. and Hoefler, T., 2021. Deep learning for post-processing ensemble weather forecasts. Philosophical Transactions of the Royal Society A, 379(2194), p.20200092.

[11] Leinonen, J., Nerini, D. and Berne, A., 2020. Stochastic Super-Resolution for Downscaling Time-Evolving Atmospheric Fields With a Generative Adversarial Network. IEEE Transactions on Geoscience and Remote Sensing.

[12] Tuck, J. and Boyd, S., 2021. Fitting Laplacian regularized stratified Gaussian models. Optimization and Engineering, pp.1-21.

[13] Cardona, J.L., Howland, M.F. and Dabiri, J.O., 2019. Seeing the wind: Visual wind speed prediction with a coupled convolutional and recurrent neural network. arXiv preprint arXiv:1905.13290.