Weather Predictive System using Machine Learning Algorithms

V Sasikala
Department of Information Technology, Vignan's Nirula Institute of Technology and Science for Women, Guntur, Andhra Pradesh, India
Email- kala.sasiv88@gmail.com

Venkata Krishna Kishore K
Department of Computer Science & Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh, India
Email- kishorekvk_1@yahoo.com

Venkatramaphanikumar S
Department of Computer Science & Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh, India
Email- svrphanikumar@yahoo.com

Pedapaga Akash Babu
Department of Computer Science & Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh, India
Email- akashpedapaga@gmail.com

M Eswar Kumar
Department of Computer Science & Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh, India
Email- eswarnaidumandela@gmail.com

N Gopi Krishna
Department of Computer Science & Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh, India
Email- nangineni9959@gmail.com

Abstract- Weather Prediction has been one of the rapidly emerging technology for ages. The Weather Prediction is a method to forecast the situations and climatic conditions by using various attributes like previous data, scientific knowledge, and experimental observations on the weather over a long time. There have been different weather prediction models evolving for ages. Each time a new model that overcomes the disadvantages of the existing model has been built. The objective of building a new model is to create a perfect and cost-efficient weather prediction model to preserve the growing expenses. This paper compares various techniques and models that have been used for Weather prediction. According to the survey, the various methods and models used for weather prediction are Bayesian Model, Fuzzy Neural Networks, Hadoop based model, Kalman filter, WRF- ARF Model, Support vector Machine, Genetic algorithms and different Machine learning, Deep learning, Soft Computing techniques and discusses the Datamining classifiers use in weather prediction focusing on Linear Regression. The working of the classifiers is demonstrated on a weatherdata.csv dataset in Weka Tool.

Keywords – Weather Forecasting, Predictive Analytics, Random Forest, Naïve Bayes, Machine Learning

I. INTRODUCTION

The Weather prediction system has very important in our daily life. It is employed in various fields like agriculture, industries, hydrology, sports, etc. It is the prediction of weather depending on the previous datasets. The weather prediction seeks a lot of attention nowadays because of its importance. The prediction of weather is not that easy due to its limitations and challenges that affect the accuracy of the prediction. After considering the dataset, it has to undergo specific pre-processing techniques like data cleaning, data integration, data reduction, and Data transformation. Data cleaning includes filling up the missing values. It is done either manually or with global constants or by the average of the remaining costs. Some systems follow the method of filling the missing values by replacing the missing values with the average value of the class attribute. In some cases, the missing values are replaced by the most repeated value, i.e., Mode. After handling the missing data's missing values, the data is cleaned by removing the unwanted values, which is also called the Noise. The most effective way to handle Noisy data is by Binning. The data is binned either by Means or Medians or by boundaries. Some other systems even follow the regression or Outlier Analysis to handle the Noise present in the data. Data integration is merging the data or collection of data at a single side collected from multiple repositories. Though there are certain issues like Tuple Duplication, Data redundancy, and Entity identification problems that arise in the data integration, the issues can be handled easily by various techniques like Chi square test and Correlation analysis.

After the Pre-processing of the data, the model is built. The model is then trained by a training data set that is specifically used to train the data. After the model is trained, the model is applied with a classifier that classifies the data from a user-specified Test dataset. Thus, the classification of data is done. The above pre-processing and classification are common to all the data mining models. It varies on the train set and test set given to the model. In the proposed methodology, the model is presented with a weather dataset as the aim is Weather Prediction. The dataset description is given in the following passages. There are many models and techniques to predict the weather, and each model has different accuracies. This paper mainly focuses on weather prediction using various classification techniques. The rest of the paper Literature Survey is presented in Section II. Proposed algorithms are explained in section III. Experimental results are presented in section IV. Concluding remarks are given in section V.

II. LITERATURE SURVEY

Nitin Singh et al. [1] proposed a technique of using machine learning algorithms to anticipate the weather. The implemented system use machine learning algorithms and data analytics like the Random forest algorithm to predict the weather. In the proposed system, the application that gets the real-time data from humidity-temperature and pressure from the sensors to predict the possibility of rain has been developed on Raspberry Pi 3B. The GPIO pins of the Raspberry Pi 3 B works between the sensors and the application. Figure 2 shows the overall idea of using Machine learning algorithms in Weather prediction [1]. The application takes and collects the data through the GPIO pins. The model is being trained in accordance with the Random Forest Algorithm. The result is generated by the system in the form of Binary digits (1/0). The result 1 indicates that there would be rainfall, and 0 indicates that there are no chances for rainfall. The resultant confusion matrix has generated 1491 correct predictions that there would be no rainfall and 122 correct predictions that there will be rainfall, which indicates that 1613 correct anticipations have been brought out of 1835 test data. Afan Galih Salman et al. [2] has come up with a technique of utilizing Deep Learning techniques for weather anticipation and prediction. Various prediction models and their prediction performances have been discussed in this paper. Those forecast models incorporate Recurrence Neural System (RNN), Contingent Confined Boltzmann Machine (CRBM), and Convolutional System (CN) models. The investigation inferred that the Contingent Confined Boltzmann Machine (CRBM) and Convolutional Neural System (CNN) could be used to predict the weather on a concourse of time-series troubles and the Recurrence Neural Network (RNN) which is implemented using heuristically optimization method can be used to predict the rainfall with fair rainfall level.

Wenying Zhang et al. [3] Initiated a strategy of anticipating climate with multiclass Bolster Vector Machines in the Issue Identification of Photovoltaic system. This is an incredible method for planning and predicting the weather depending upon Photovoltaic power and meteorological information. Using this technique, weather kinds were assumed by using data analytics instead of weather instruments. The Direct Support vector machine has provided better fault detection. The support vector machine has established a model for weather prediction for training multiclass predictors. The classified results depend upon the kernel. i.e., the type, parameters, and the soft margin coefficient through the Support Vector Machine are capable of classification. This technique has come up with a technique of using the Particle Swarm Optimization (PSO) algorithm to maximize the parameters of the kernel. The technique was proven to be effective and feasible by the results. Y Beeharry et al. [4] has come up with another idea of implementing a fully encrypted end-to-end system from actual time info collection to actual time climate prediction. The prediction algorithm is implemented by using the Devops insight together with an android mobile application. The NoSQL database service is used to implement the storage on the International Business machine Cloud platforms.

In this method, an Arduino based microcontroller node was setup with all the required sensors to obtain the data. The block diagram of this proposed methodology is shown in the figure 3.

The IBM cloud database which stores the data in the document format is connected to the node through an internet gateway. The K-NN algorithm was implemented on the Cloud platform using JAVA programming language. The Android application that forecasts the weather for the next hour or next two hours has been developed and it needs to be installed on the user's android mobile phone to predict the weather. When the system was organized and arranged on the International Business Machine (IBM) cloud platform followed by the actual time storage systems, it was able to demonstrate the rapid real time prediction. Christy Kunjumon et al. [5] proposed a technique to predict the weather at a particular place by making use of scientific knowledge and Data mining techniques. The method includes usage of various data mining techniques including Supervised and Unsupervised algorithms, Naïve- Baye's Algorithm, Decision tree classification algorithm, k- Medoids algorithm, Hadoop map reduces, Support vector machine, Artificial neural networks and FP Growth algorithm. The ANN is an information processing model that consists of artificial neurons. This model lets the flow of information across the network during the training phase. The functioning of the Human brain is replicated in the model. The Support vector machine caters a better assortment that generates more difficult and complex boundaries between classes since it dependent on the Statistical learning theory. The SVM separates various classes by finding out the optimal hyper plane in linearly separable classes. The required Support vectors lie on the margin of the hyper plane that is found out by the SVM. Frequent Pattern Growth algorithm is one of the scalable techniques that is used to predict the weather. It generates frequent patterns that are used in pattern mining to forecast the possible patterns of weather. The K-Medoids algorithms is one of the clustering algorithms and it is similar or related to medoid shift. This algorithm focuses on finding out the representative objects that are also known as medoids in the groups or Clusters. This algorithm optimizes the distances between the points that are labelled in the clusters and the center of the clusters. This algorithm works with an arbitrary metrics of distances between data points and it chooses the centers from the data points.

Naïve Bayesian methodology for arrangement is a measurable and straight classifier which predicts class name for information case based on dispersion of property estimations. This is a parametric arrangement where the size of classifier stays fixed. Liviu Oan˘a et al.,[6] has come up with another technique of using genetic algorithms for Numerical weather prediction systems. The main intent of this study is to improve the forecast of 2-meter temperature as well as relative humidity and to optimize the physical parameterization configuration. The parameterization sub modules include Microphysical parameterization, Radiation parameterization, Convective parameterization, Soil and soil surface processes that can estimate soil physics, Planetary boundary layer physics and Urban physics. In this approach a new genetic algorithm named WRF-GA is written. The physical arrangement choices like microphysics, cumulus material science, longwave radiation material science, shortwave radiation material science, surface physical science, PBL material science and soil material science are encoded on the chromosomes. 14 Chromosomes were combined into a generation. A 110x70 grid of 10 -kilometer resolution was used for the experimentation. The evaluation was done from 5 weather stations basing on an absolute average error. Based on forecast error, the chromosomes were sorted in an ascending order. The first chromosome will become snobbish. The top 11 chromosomes are selected for the random crossover operation. They will repeat in pairs each pair creating two new off springs. This procedure is rehashed until 30^{th} era. The optimization performed well in finding ideal physical designs for mugginess expectation. In any case, demonstrated poor outcomes for temperature figure. More examinations should be directed to have an away from over the utility of utilizing Hereditary Algorithmic strategies for physical parameterization enhancement. Sunil Navadia et al.,[7] proposed a strategy to gauge the odds of precipitation by utilizing prescient investigation in Hadoop. The proposed framework fills in as an apparatus that takes in the precipitation information from a lot of information as info and predicts the future precipitation with min, max and normal precipitation proficiently. Prescient expository models catch connections among numerous components in an informational collection to survey chance with a specific arrangement of conditions to relegate a score or a weight. Considering these loads the expectation occurs. We will utilize Guide decrease and Pig orders to break down the informational collections and to perform different procedure on the informational collection. In view of the earlier year's recorded climate informational collection we can anticipate the future climate. Tempest can be anticipated by utilizing the earlier year's informational collection. The information is ordered utilizing the Help Vector Machine (SVM). Utilizing this it can anticipate most extreme Downpour Tempest. Guide Lessen System is utilized for the Downpour Tempest Forecast.

Apache pig is one of the platforms of the Hadoop that creates programs. Guide Decrease is a system utilized for effectively composing applications that procedure enormous measures of information on huge groups, issue open minded way. By using Hadoop, the prediction happened in the first level, but some errors will be removed by using a Naïve Bayes Algorithm in the next stage. The prediction and anticipating of earthquakes, floods can be done by utilizing Naïve Bayes Algorithm this is the future scope of the project. Kan Li et al.,[8] has come up with a technique

of using A unpleasant set-based fluffy neural system calculation in climate forecast models. To abstain from showing up at a nearby least worth, the least square calculation (LSA) is utilized in the learning procedure of a fluffy neural system to get worldwide intermingling. Since the structure of fluffy neural systems, the quantities of rules and the underlying loads are hard to be resolved, here the unpleasant sets technique is acquainted with choose the quantities of rules and unique loads. One should have good knowledge about the 4 layers of the fuzzy neural networks for the model to be initiated. The harsh set strategy is likewise used to choose the underlying end parameters. Unpleasant Set hypothesis is an apparatus to handle fluffy and questionable information. A data framework S={U,A,V ,f}, where U is a limited arrangement of items, U={x_1,x_2 … , x_n}, A is a limited arrangement of traits, The qualities in An are additionally grouped into two disjoint subsets: Condition characteristics C and choice properties D.

$$A = C\ D.$$

$$V = V \cup 1\,V \cup 2 \ldots Vp(p \in A)$$

Vp is a domain of attribute p. $f: U\,A \rightarrow V$ is a total function such that f(x ∪ × i, q) $V \in q$ for every $q \in A$, x_i∈U. The learning algorithm is used to train the fuzzy neural network. Here the least square algorithm is used to decide the parameters. Finally, we conclude that this model is used to predict the weather on paper. The numbers of fuzzy rules decide the numbers of the hidden layer and initial weights decide the initial values of the conclusion parameters. Aye Nandar et al.,[9] lodged another method of using Bayesian systems, or conviction systems, show restrictive likelihood and causality connections between factors. This work applies BN to show the spatial conditions among the diverse meteorological factors for climate (precipitation and temperature) forecast over Myanmar. In this work, local and worldwide climate information that are adding to precipitation expectation of Myanmar are utilized for precipitation forecast. In this historical data of various weather stations from 1990 to 2006 were used to predict the weather. A Bayesian system, conviction arrange or coordinated non-cyclic graphical model is a probabilistic graphical model that speaks to a lot of irregular factors and their contingent independencies by means of a coordinated non-cyclic chart (DAG). This model needs data of various factors like East India Sea Surface Temperature, Southern Oscillation Index, monthly precipitation of previous year, precipitation of previous month, temperature, precipitation amount. Historical data is divided into testing data and training data. The different phases in the model are data collection and preparation, predictor analysis process, BN training, prediction process, Evaluate predicted result. We conclude that this model can give acceptable accuracy in terms of experimental results.

A.K. Pandey et al.,[10] constructed a Hadoop based Climate Forecast Model to arrange the Climate Information. In this model, the climate expectation information is created from different sources, for example, radar, ships, ground perception, and so on. It contains certain helpful and futile data for the expectation of climate information and as unstructured information. Further, in this work, the Hadoop system is applied to process this unstructured information. Word tally calculation is utilized to locate the general state of the day. Further fluffy rationale (FL) and fake neural system fluffy interface framework (ANFIS) strategies are explored for exact forecast of climate information dependent on mean square mistake. The test results show that this strategy gives exact outcomes contrasted with different strategies. The means associated with this model are information assortment, pre-handling of information utilizing word check calculation, grouping of qualities is recognized, expectation utilizing the ANFIS and Fluffy Rationale. Dataset has five traits for the expectation of climate information. Four traits are temperature, precipitation, stickiness and ocean level. The fifth quality is the class characteristic which shows the expectation of climate information into eight distinct classes. These classes are smoke, murkiness, haze, fog, downpour, shady, clear and residue. In this model we have utilized two information mining apparatuses for example ANFIS and FL. ANFIS strategy predicts climate information more precisely is finished up from the outcomes got. Future extension is that other delicate figuring procedures are examined for better expectation of climate information. Nikolajs Bogdanovs et al.,[11] designed an algorithm that is in light of the authentic information utilizing a Kalman Channel. This calculation impressively builds the precision of the momentary gauge of outer air temperature. This calculation adjusting anticipated air temperature inside the following three hours is thought of. The handling of temperature information utilizing Kalman Channel gives a diminishing in anticipated temperature mistakes. We use python in this model for the usage of the exactness improving calculation of anticipated temperatures. Kalman channel-based calculation is normally intended to wipe out precise mistakes and effectively incorporates gauge information dependent on the NWP model, which significantly rearranges the centre of the calculation. It diminishes the pre-handling time. The fundamental components of the KF calculation are framework state vector, Xk; mistake covariance grid, Pk; framework commotion covariance lattice, Qk; framework model, Φk; estimation vector, Zk; estimation clamour covariance network, Rk; and estimation model Hk. This calculation can altogether build the exactness of the momentary figure of outside air temperature. The considered calculation gives temperature gauge adjustment to the following three hours. The temperature information handling by KF gives predisposition (decrease100 times) and standard deviation (decline by 20%) diminishing estimated temperature blunder. The future work incorporates presenting of versatile tuning of KF clamour frameworks

coefficients makes quicker set up of information handling. As climate is viewed as a nonlinear powerful framework, another sort of KF – unscented Kalman channel - can be applied for better handling with framework nonlinearity.

III. PROPOSED METHOD

The performance evaluation of each classifier is discussed in a tabular manner and the classifier with best accuracy discovered. All the weather prediction techniques are demonstrated by applying the classifiers on a dataset namely weatherdata.csv in the Weka tool. The dataset consists of 32686 Values with 10 attributes naming temperature, heat_index, humidity, pressure, wind direction, wind speed, precipitation, gust speed, sea level pressure, conditions. The conditions attribute acts as a class label. There are 28 attributes in the dataset. Those classes are Smoke, Clear, Haze, Overcast, Scattered Clouds, Shallow Fog, Mostly Cloudy, Fog, Partly Cloudy, Fog Patches, Thunderstorms with Rain, Rain, Light Rain, Light Drizzle, Drizzle, Mist, Volcanic Ash, Thunderstorm, Light Thunderstorms with Rain, Light Thunderstorm, Squalls, Heavy Rain, Light Haze, Sandstorm, Widespread Dust, Funnel Cloud, Heavy Thunderstorms with Rain, Heavy Thunderstorms with Hail, Light Rain Showers. The steps followed in the weather prediction system are depicted as a Flow Chart in the Figure 1.

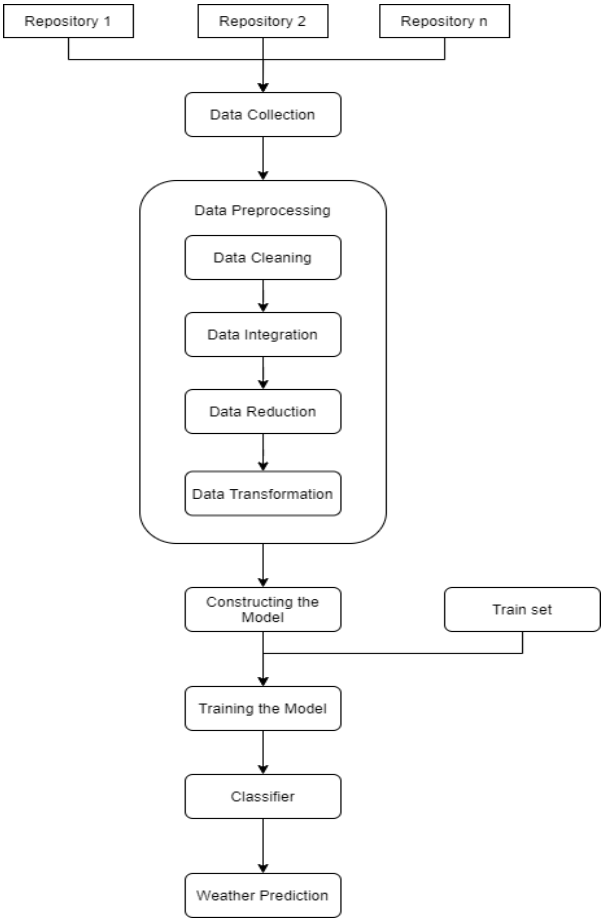

Fig. 1: Flowchart of Weather Prediction System

Weather Prediction can be done by using various classifiers in data mining techniques. Some of the techniques that assist in Weather Prediction are discussed in the following. Each technique has its own advantages and disadvantages and each of the classification technique becomes handy depending upon the requirement and the conditions.

Naïve Bayes: This Naïve Bayes classifier depends on easiest Bayesian system models. This classifier works on Bayes theorem. It predicts the probabilities for each record to have membership in a class. This classifier is exceptionally versatile requiring various parameters in an issue. It is based on conditional probability and the attributes however independent with each other. The class with highest probability is known as Maximum A Posteriori (MAP). The Naïve Bayes classifier works on the following formula.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Where, $P(c|x)$ is the Posterior Probability, $P(x|c)$ is the Likelihood of the class, $P(c)$ is the Class Prior Probability and $P(x)$is the Predictor prior probability The overall equation is as follows

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \ldots \times P(x_n|c) \times P(c)$$

KNN Classifier: K-Nearest Neighbours is an algorithm for classification and regression. A data object is grouped utilizing the lion's share vote of its closest neighbours. For instance, on the off chance that K=1, at that point the item is appointed to class of the single closest neighbor. In order to find out how near a data object is to a neighbor and finally assign it a class, there are several distance measures such as Euclidean distance, Manhattan Distance, Minkowski Distance.

Linear Regression: In statistics, linear regression is a direct way to deal with displaying the connection between a scalar and at least one illustrative. On the other hand, the least squares approach can be utilized to fit models that are not direct models. In this manner, even though the expressions "least squares" and "straight model" are firmly connected, they are not equivalent.

The Linear Regression: The equation has the form

$$Y = a + b * X$$

Where Y is the dependent variable, X is the independent variable, b is the slope of the line and a is the y-intercept. The general form of the Linear Regression equation is

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

Weather Forecasting is a basic zone of investigation in regular day to day existence. Climate for future is one of the most significant ascribes to gauge since agribusiness areas, just as numerous ventures, are to a great extent reliant on the climate conditions.

The process for using regression to predict climatic changes is as follows

1. Research the branch of knowledge so you can expand on crafted by others.
2. Collect information for the applicable factors.
3. Specify and survey your relapse models.
4. If you have a model that sufficiently fits the information, use it to make expectations.

For a decent regression model, you need to incorporate the factors that you are explicitly trying alongside different factors that influence the reaction to keep away from one-sided results. Cross-approval decides how well your model sums up to other informational indexes by dividing your information. Weather prediction are made by gathering however much information as could reasonably be expected about the present condition of the air (especially the temperature, stickiness and wind) and utilizing comprehension of environmental procedures (through meteorology) to decide how the air advances later. Based on these factors the temperature can be predicted and with addition to some attributes (wind direction, radiation and wind speed) we can predict the weather.

Decision Tree: Decision tree is a calculation that utilizes a tree like chart or model of Decisions and their potential results to foresee an official Decision, this calculation utilizes contingent control articulation. A Decision tree is a calculation for moving toward discrete-esteemed objective capacities, in which Decision tree is signified by an educated capacity. For inductive learning these sorts of calculations are popular and have been effectively applied to abroad scope of errands. We offer name to another exchange that is whether it is genuine or misrepresentation for which class name is obscure and afterward exchange esteem is tried against the Decision tree, and after that from root hub to yield/class name for that exchange a way is followed. Decision guidelines decides the result of the substance of leaf hub. As the rules have the type of 'On the off chance that condition 1 and condition 2 yet not condition 3, at that point result. Decision tree assists with deciding the most exceedingly terrible, best and anticipated qualities for various situations, rearranged to comprehend and decipher and permits expansion of new potential situations.

J48 Decision Tree: The J48 Decision tree is generated by the C4.5 algorithm. C4.5 fabricates decision trees from a training set of data similarly as ID3, utilizing the idea of information entropy. The training set is a set

{\displaystyle S = {s_{1}, s_{2}, \ldots }}S = {s_{1}, s_{2}, \ldots }

of effectively grouped examples. Each example {\displaystyle s_{i}}s_{i} comprises of a p-dimensional vector

{\displaystyle \left(x_{1,i}, x_{2,i}, \ldots, x_{p,i}\right)} (x_{{1, i}}, x_{{2, i}}, \ldots, x_{{p, i}})

Where,

{\displaystyle x_{j}}x_{j} speak to trait esteems or highlight features of the sample.

At every node, the algorithm picks the data attribute that most viably parts its arrangement of data samples into subsets advanced in one class or the other. The parting rule is the standardized information gain (contrast in entropy). The

data attribute with the most elevated and standardized information gain is picked to settle on the decision. The C4.5 algorithm then recurses on the apportioned sub lists.

Filtered Classifier: Most of the times, a filter needs to be used on the data before a Classifier is developed. The filter is applied on the data to perform various operations like Removal of useless attributes, Transformation, Discretization, Addition of attributes, Normalization of instances etc. Usually Filtering in Weka is done by using the Filter Class in Weka. Instead, Weka has a class named Filtered Classifier. This class runs arbitrary classifier on data that has been passed through an arbitrary filter[13]. Like the classifier, the structure of the filter depends only on the training data and test cases will be handled by the channel without changing their structure. On the off chance that inconsistent case weights or attribute weights are available, and the filter or the classifier can't manage them, the instances as well as attributes are resampled with substitution dependent on the weights before they are passed to the filter or the classifier[14] (as suitable).

IV. EXPERIMENTAL RESULTS & DISCUSSIONS

The results from each classifier is tabulated with the attributes like Accuracy, Error rate, Kappa statistic, mean absolute error, Root mean squared error, Relative squared error, Root relative squared error and Time taken to build the model. Accuracy is the percentage measure of the correctly classified instances in the given data. Error rate is the percentage measure of the incorrectly classified instances in the given data.

$$Accuracy = \frac{TP + TN}{P + N}$$

$$Error\ Rate = 1 - Accuracy$$

Where TP is the True Positives, TN is the True Negatives, P is the Total number of Positives and N is the Total Number of Negatives.

Kappa statistic is the measure of the inter-rater reliability of the data. It is utilized to control just those occasions that may have been accurately characterized by some coincidence.

$$Kappa\ Statistic = \frac{(OA - EA)}{(1 - EA)}$$

Where OA is the Observed Accuracy, EA is the Expected Accuracy.

Mean absolute error alludes to the mean of the supreme estimations of every prediction error on all occurrences of the test dataset. Prediction error is the distinction between the real value and the anticipated value for that case. The Root mean squared error is the standard deviation of the errors that occur during the prediction on the dataset.

$$Mean\ Absolute\ Error = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$$

$$\text{Root mean squared error} = \sqrt{\frac{\sum_{i=1}^{n}|x_i - \bar{x}|^2}{n}}$$

Where x_i is the i^{th} value of x, $\bar{x}$ is the mean of x, Relative squared error is the error is made comparative with what it would have been if a basic predictor had been utilized. The basic predictor being referred to is only the average of the real values from the training data. Root relative squared error is the squared root of the relative squared error.

$$\text{Relative squared error} = \frac{\sum_{j=1}^{n}(P_{ij} - T_j)^2}{\sum_{j=1}^{n}(T_j - \bar{T})^2}$$

$$\text{Root Relative squared error} = \sqrt{\frac{\sum_{j=1}^{n}(P_{ij} - T_j)^2}{\sum_{j=1}^{n}(T_j - \bar{T})^2}}$$

Where P_{ij} is Predicted value by the model i for the record j, T_j is the Target Value for Record j, $\bar{T}$ is given by

$$\bar{T} = \frac{1}{n}\sum_{j=1}^{n} T_j$$

Table 1: Experimental results with various classifiers

Name of the Classifier	Kappa Statistic	Mean Absolute Error	Root mean Squared Error	Relative absolute error (%)	Root relative squared error (%)	Time taken to build model (in sec)
Naïve Bayes	0.1481	0.0452	0.1508	95.1525	97.8744	0.23
Naïve Bayes Updateable	0.1481	0.0452	0.1508	95.1525	97.8744	0.07
Stacking	0	0.0475	0.1541	100	100	0.5
Attribute selected	0.3346	0.0378	0.1429	79.692	92.7583	2.37
Filtered Classifier	0.3614	0.0363	0.1431	76.4677	92.8981	0.86
Multi Scheme Classifier	0	0.0475	0.1541	100	100	0.3
Iterative Classifier Optimizer	0.1392	0.0456	0.151	96.0626	97.986	144.1
Input mapped Classifier	0	0.0475	0.1541	100	100	0.01
OneR Classifier	0.0502	0.0376	0.194	79.2269	125.9034	0.19
Decision Stump Classifier	0.0012	0.0473	0.1538	99.5989	99.8214	0.36
Hoeffding tree	0.1481	0.0452	0.1508	95.1525	97.8744	1.63
J48 Tree	0.3614	0.0363	0.1431	76.4677	92.8981	1.03
Random Tree	0.3295	0.0336	0.1614	70.7933	104.72	0.79
REP Tree	0.3392	0.037	0.1431	77.9721	92.8873	1.83

V CONCLUSION

The weather prediction has become one of the most essential entity now a days. To improve the risk management systems and to know the weather in coming days in an automatic and in scientific way, many models have been emerging to assist in weather Prediction. In this paper, eleven methods that are used in weather prediction are compared as a survey and five most used and well-known data mining classifiers are explained. In the result section, the results from 15 various classifiers and their results such as Accuracy, Error rate, Kappa statistic, mean absolute error, Root mean squared error, Relative squared error, Root relative squared error and Time taken to build the model are tabulated. The results show that the J48 Tree classifier that uses C4.5 algorithm and the Filtered classifier has output the results of high accuracy than all the other classifiers that were used. When coming to the time taken to build the model, The Filtered Classifier outperforms all the other classifiers in solving the Problem under scrutiny.

REFERENCES

[1] N. Singh, S. Chaturvedi, and S. Akhter, "Weather Forecasting Using Machine Learning Algorithm," *2019 Int. Conf. Signal Process. Commun. ICSC 2019*, pp. 171–174, 2019.

[2] A. G. Salman, B. Kanigoro, and Y. Heryadi, "Weather forecasting using deep learning techniques," *ICACSIS 2015 - 2015 Int. Conf. Adv. Comput. Sci. Inf. Syst. Proc.*, pp. 281–285, 2016.

[3] W. Zhang, H. Zhang, J. Liu, K. Li, D. Yang, and H. Tian, "Weather prediction with multiclass support vector machines in the fault detection of photovoltaic system," *IEEE/CAA J. Autom. Sin.*, vol. 4, no. 3, pp. 520–525, 2017.

[4] Y. Beeharry, T. P. Fowdur, and J. A. Sunglee, "A Cloud-Based Real-Time Weather Forecasting Application," pp. 294–297, 2019.

[5] C. Kunjumon, S. S. Nair, S. Deepa Rajan, L. Padma Suresh, and S. L. Preetha, "Survey on Weather Forecasting Using Data Mining," *Proc. IEEE Conf. Emerg. Devices Smart Syst. ICEDSS 2018*, no. March, pp. 262–264, 2018.

[6] L. Oana and A. Spataru, "Use of genetic algorithms in numerical weather prediction," *Proc. - 18th Int. Symp. Symb. Numer. Algorithms Sci. Comput. SYNASC 2016*, pp. 456–461, 2017.

[7] S. Navadia, P. Yadav, and J. Thomas, "Measuring and Analyzing Weather Data," pp. 414–417, 2017.

[8] K. Li and Y. S. Liu, "A rough set based fuzzy neural network algorithm for weather prediction," *2005 Int. Conf. Mach. Learn. Cybern. ICMLC 2005*, no. August, pp. 1888–1892, 2005.

[9] A. Nandar, "Bayesian network probability model for weather prediction," *Proc. 2009 Int. Conf. Curr. Trends*

Inf. Technol. CTIT 2009, pp. 120–124, 2009.

[10] A. K. Pandey, C. P. Agrawal, and M. Agrawal, "A hadoop based weather prediction model for classification of weather data," *Proc. 2017 2nd IEEE Int. Conf. Electr. Comput. Commun. Technol. ICECCT 2017*, pp. 1–5, 2017.

[11] N. Bogdanovs, V. Bistrovs, E. Petersons, A. Ipatovs, and R. Belinskis, "Weather Prediction Algorithm Based on Historical Data Using Kalman Filter," *Proc. - 2018 Adv. Wirel. Opt. Commun. RTUWO 2018*, pp. 94–99, 2018.

[12] "Decision tree for the Weather dataset | Download Scientific Diagram." [Online]. Available: https://www.researchgate.net/figure/Decision-tree-for-the-Weather-dataset_fig1_232172450.

[13] "Classifying unseen test data with a filtered classifier - Java Data Science Cookbook." [Online]. Available: https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781787122536/4/ch04lvl1sec46/classifying-unseen-test-data-with-a-filtered-classifier.

[14] Edara, D.C., Vanukuri, L.P., Sistla, V. *et al.* Sentiment analysis and text categorization of cancer medical records with LSTM. *J Ambient Intell Human Comput* (2019). https://doi.org/10.1007/s12652-019-01399-8