iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** VI **Month of publication:** June 2019

DOI: http://doi.org/10.22214/ijraset.2019.6373

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

An Integrated Approach for Weather Predictions using Machine Learning Technique

Manjushri U A[1], Mallikarjuna S B[2]
[1] *Student, Dept. of Computer Science and Engineering, BIET college, Karnataka, India.*
[2] *Assistant Professor, Dept. of Computer Science and Engineering, BIET college, Karnataka, India.*

Abstract: ***The weather plays a very important role in our daily activities. Time makes a big difference in our daily lives and is often an unconscious and even conscious factor in making many decisions. Weather forecasts are therefore excessively used in many areas such as food security, agriculture and science. Traditionally, weather forecasts were made using physical simulations by solving the fluid dynamics and thermodynamics equations. However, the traditional system is uncertain due to incomplete understanding of complex atmospheric processes that limit accurate weather forecasts for the next few days. Therefore, the proposed application tends to inculcate the decision tree using the machine learning technique, which provides approximate and close results predicting the weather for the next 5 days.***
Keywords: Decision tree, ID3 Algorithm, Sliding window operation, Weather Predictions.

I. INTRODUCTION

A. Overview

Weather plays an important role in the everyday routine of a layperson or businessman, forcing a person to know future weather conditions. Climate change can have serious consequences for the availability of human resources, especially in countries. Climate change will have a significant impact on public health. Climate change refers to long-term changes in weather patterns and extreme weather patterns. This can lead to changes in the threat to human health and multiply existing health problems. The time that is today is the problems that stimulate the world. The application of weather forecasts is an art of science and technology used to provide the state of the atmosphere of a place. The weather forecasting system has been useful in preventing damage to life and large areas of property. Weather forecasting is a field of meteorology that collects dynamic data related to current weather conditions such as fog, precipitation, temperature, wind, etc. Weather forecasts can be used to plan activities and events to plan further and survive.Meteorology department used to update the knowledge of the current state of the atmosphere by,

1) Satellites measures radiation from earth's plane and the impression.
2) Balloons and aircraft measures the bit of the air that they are passing through.
3) Buoys and land stations measures the lower part of the atmosphere.
4) Radar systems measures the indication of emission from rain drops and snowflakes.

Data collected in different states is distorted into a digital representation of recent atmospheric conditions. This process is called assimilation. Small changes in atmospheric conditions result in very different weather conditions, so it is essential that the current state of the atmosphere is represented as accurately as possible. Climatic variations need to be addressed and analysis needs to be done to help farmers maximize the productivity of their crops.

B. Existing System

Previously, different methods were used for weather forecasts, some of which were based on intuition, while others were based on scientific and mathematical bases. Digital Weather Forecasting is a modern weather forecast based on the fact that changes in the atmosphere obey many known physical principles and that the same laws can be used to predict the future state of the atmosphere based on current conditions. Traditionally, weather forecasts have always been made by physically simulating the atmosphere as a fluid. The current state of the atmosphere is sampled. The future state of the atmosphere is calculated by solving the numerical equations of thermodynamics and fluid dynamics.

Current determining innovation includes reenactments in light of material science and differential conditions. Numerous new methodologies utilized neural network systems for weather predictions. While others used models which had a probabilistic approach, for example, bayesian systems to predict the weather. Another approach concentrated only on a more particular instance of anticipating extreme climate for a particular topographical area.

1) *Disadvantages:* This traditional system of differential equations governing the physical model is sometimes unstable in the presence of disturbances and uncertainties when measuring the initial conditions of the atmosphere. This leads to an incomplete understanding of atmospheric processes.

C. Problem Statement

Nowadays, people used to experience different weather conditions in different seasons because of weather changes, pollution and their reactions. So the main objective is to develop a soft computing system that provides weather predictions for next 5 days using machine learning techniques.

D. Proposed System

The proposed system uses the existing Google API dataset to extract the necessary meteorological information. The proposed system uses a decision tree and conditions using the id3 algorithm, which provides approximate and nearby results to predict the weather for the next 5 days. The proposed system ranks the weather in four groups: sunny, cloudy, humid and rainy.

1) *Advantages*
1) ID3 algorithm handles each attribute with different cost.
2) Missing attribute values in training data set are not used in information gain and entropy calculations.
3) It creates a threshold and then splits the list into those whose attribute value is above the threshold and those that are less than or equal toit.
4) ID3 algorithm goes back through the tree once it has been created, and attempts to remove branches that are not needed, by replacing them with leaf nodes.

II. LITERATURE SURVEY

The following are some of the works that were carried out by specific persons on Weather Predictions:

A. In 1922, English scientist Lewis Fry Richardson published "Weather Prediction By Numerical Process", after finding notes and derivations he worked on as an ambulance driver in World War I. He described therein how small terms in the prognostic fluid dynamics equations governing atmospheric flow could be neglected, and a finite differencing scheme in time and space could be devised, to allow numerical prediction solutions to be found.

B. In 1940's, for weather forecasting, used the techniques of Numerical Weather Prediction (NWP). This method is based on the fact that gases of the atmosphere follow a number of physical principles. If the current conditions of the atmosphere are known, these physical laws may be used to forecast the future weather. Since the late 1940s, there has been a steady growth as regards the use of mathematical models in weather forecasting. These procedures have been made possible because of advancement in the formulation of mathematical models. A series of mathematical equations is used to develop theoretical models of the general circulation of the atmosphere. These equations are also used to specify changes in the atmosphere as the time passes on. For these equations certain weather elements like air movements, temperatures, humidity, evaporation at the ground, clouds, rain, snow and interactions of air with ground and oceans are taken into account.

C. In 2010, Jethangir and Onaiza proposed learning vector Quantization (LVQ) that is used for monsoon rainfall prediction. 45 years monsoon rainfall data is used to train Neural Network and evaluate the performance of these models over a test period of 5 years from 2005-2009. The results were compared with multiple linear regressions and statistical downscaling models, but the results reveals neural network has better performance in terms of accuracy, and also in terms of greater lead time and fever required resources.

D. In 2011, Kesheng and Lingzhi presented a novel modular type support vector machine to simulate weather predictions. Firstly, a bagging sampling technique is used to generate unlike training sets. Secondly, changed kernel function of SVM with base models and then trained to formulate different regression based on the poles apart training sets. Thirdly, the partial least square (PLS) technology is used to select decide the appropriate number of SVR combination member. Finally, a V-SVM was produced by learning from all base models.

E. In 2013, Mark, Bobby, Yung and Beth proposed time series analysis is used as prediction algorithm. Decision support system for Agriculture management using prediction algorithm aimed to develop a system that will determine the trend of rainfall and evaporation using time series analysis as its prediction algorithm, to develop web- based application that displays graphs and tables according to the result of the prediction algorithm, and to utilize a classification of crops that aids farmers as basis for recommendation according to the predicted amount of rainfall per quarter. The system is found useful in terms of efficiency, reliability.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

III. THEORETICAL BACKGROUND

The weather prediction application involves two major components: the Sliding window operation and ID3 algorithm.

A. Sliding window Operation

The sliding window algorithm divides a collection of data into groups. More specifically, the sliding window algorithm is used to segment a collection of historical weather data into week long windows. The purpose of separating the historical 30 days data is to align the data trends with the current week's weather.

The goal of this segmentation is to find a closely related trend of past data approximate to the current date to predict the upcoming days weather.

Weekly weather trends may not align perfectly with historical data. There may be varying conditions that could offset a historical trend that is relatively close to current conditions.

The Sliding Window Algorithm accounts for this potential offset by sampling from single day data set. By dividing up this sample data, the algorithm can determine the best fit trend and predict the following day's weather.

B. ID3 Algorithm

ID3 (Dichotomiser Iterative 3) is an algorithm invented by Ross Quinlan that is used to generate a decision tree from a dataset. Here we used the ID3 algorithm that uses the information gain measure to select among the candidate attributes at each stage of tree growth.

The information gain is simply the expected reduction in entropy caused by partitioning the datasets based on their attributes. The ID3 algorithm is a descending construct of the decision tree by recursively selecting the "best attribute" to use on the current node in tree.Once is selected for the current attribute,

it generates child nodes, one for each possible value of the selected attribute. The algorithm partitions the dataset using the possible value of their attribute and assigns these subsets to an appropriate child node. The above process is repeated for each child node until all data sets associated with a node are all positive or all negative.

1) Information Gain, H(S, A) = Entropy(S) $- \sum_{v \in Values\,(A)} \frac{|Sv|}{|S|} Entropy\,(Sv)$

2) Entropy (S) = $\sum_{i=1}^{n} P_i\, log_2 P_i$

where,

S is the past dataset of weather related data.

A is the attributes of weather related data and

P_i is the proportion of S belonging to class i.

IV. SYSTEM REQUIREMENTS

A. Hardware Requirements

1) System : Pentium IV 2.4 GHz
2) Hard Disk : 500 GB
3) Ram : 4 GB
4) Keyboard : Standard keyboard

B. Software Requirements

1) OS : Windows XP/7 or higher versions
2) IDE : Android Studio 2.2.3
3) Coding : Java
4) Device : AVD for Android SDK.
5) Version : Android SDK 2.2 to 4.1.
6) Database : MySQL.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

V. SYSTEM ANALYSIS

A. System Architecture

The proposed system will follow three-tier architecture. In three-tier architecture, application will run as the traditional client/server model but from the web server.

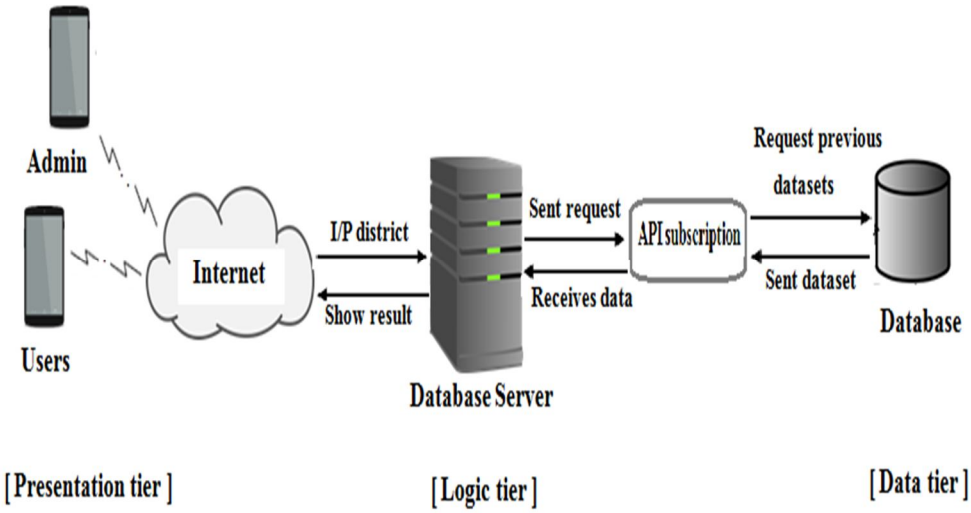


Fig. 1 System architecture

1) *Client/Presentation Tier:* This tier includes all the content or forms to be displayed for the users. It is the form which provides the interface to end user. Programmer uses this tier to get or set the data back and forth.
2) *Logical Tier:* In the logical tier, the actual processing of the data and the logic behind the implementation of the application will be present. This tier contains classes, which can be used to write the functions, and also works as a mediator between the presentation tier and data tiers.
3) *Data Tier:* Data Tier contains methods and classes that deal with passing and storing data to the data storage layer. Queries or stored procedures are used to access the data from the database or to perform any operation to the database.

VI. IMPLEMENTATION

A. Modules

1) *Registration Activity:* This module is used to implement registration page for the users to register with the necessary information.
2) *Users Login Activity:* This module provides the login page for registered users to login securely with their username and password through the android application.
3) *Admin Login Activity*: This module helps the admin to see the database that contains users information, users logs and results of weather predictions.
4) *Location Setting Activity*: This module is used to give input as the district of karnataka to access weather predictions for 5 days.

B. ID3 Algorithm

Following are the steps to create decision tree using ID3 algorithm,

1) Previous information about weather data from google API is taken as input, S.
2) Create root node for the tree by selecting the attribute that has maximum Information Gain (S,A).
3) If all examples are positive, return leaf node ‘positive’, else if all examples are negative, return leaf node ‘negative’.
4) Calculate the entropy of current state, Entropy (S) considering all input dataset.
5) For each attribute, calculate the entropy with respect to the attribute ‘x’ denoted by H(S, x).
6) Select the attribute which has maximum value of Information Gain(S, x) as node.
7) Remove the attribute that offers highest Information Gain from the set of attributes.
8) Repeat until we run out of all attributes, or the decision tree has all leaf nodes.

VII. RESULTS

The following snapshots define the results or outputs that are obtained after step by step execution of all modules of the system.

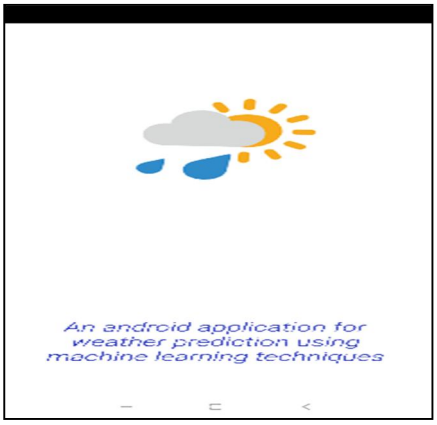


Fig 2: Splash Screen

The Fig 2 shows the splash screen that appears for 3 seconds before entering the android application.

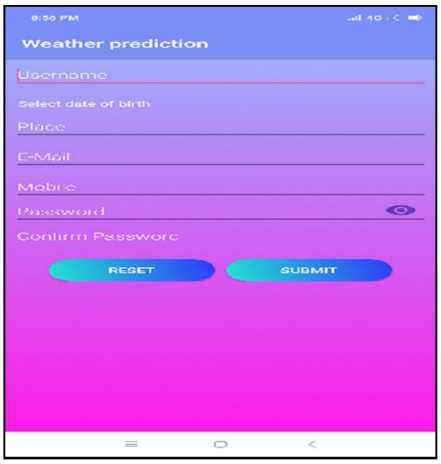


Fig 3: Registration Page

The Fig 3 shows the registration page for users to create their account by entering the personal details.

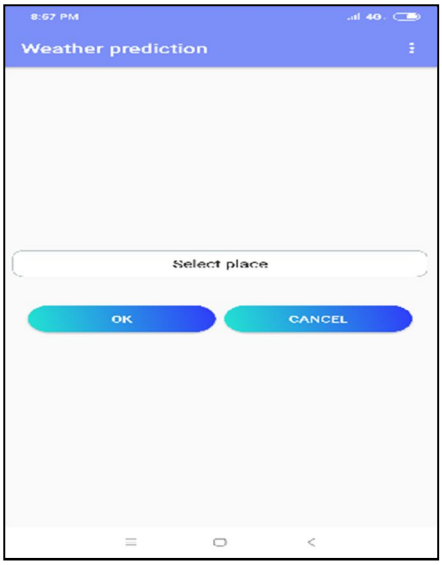


Fig 4: Location selecting page

The Fig 4 shows the Location selecting page to access weather predictions for selected district of Karnataka.

Fig 5: Weather prediction for next day

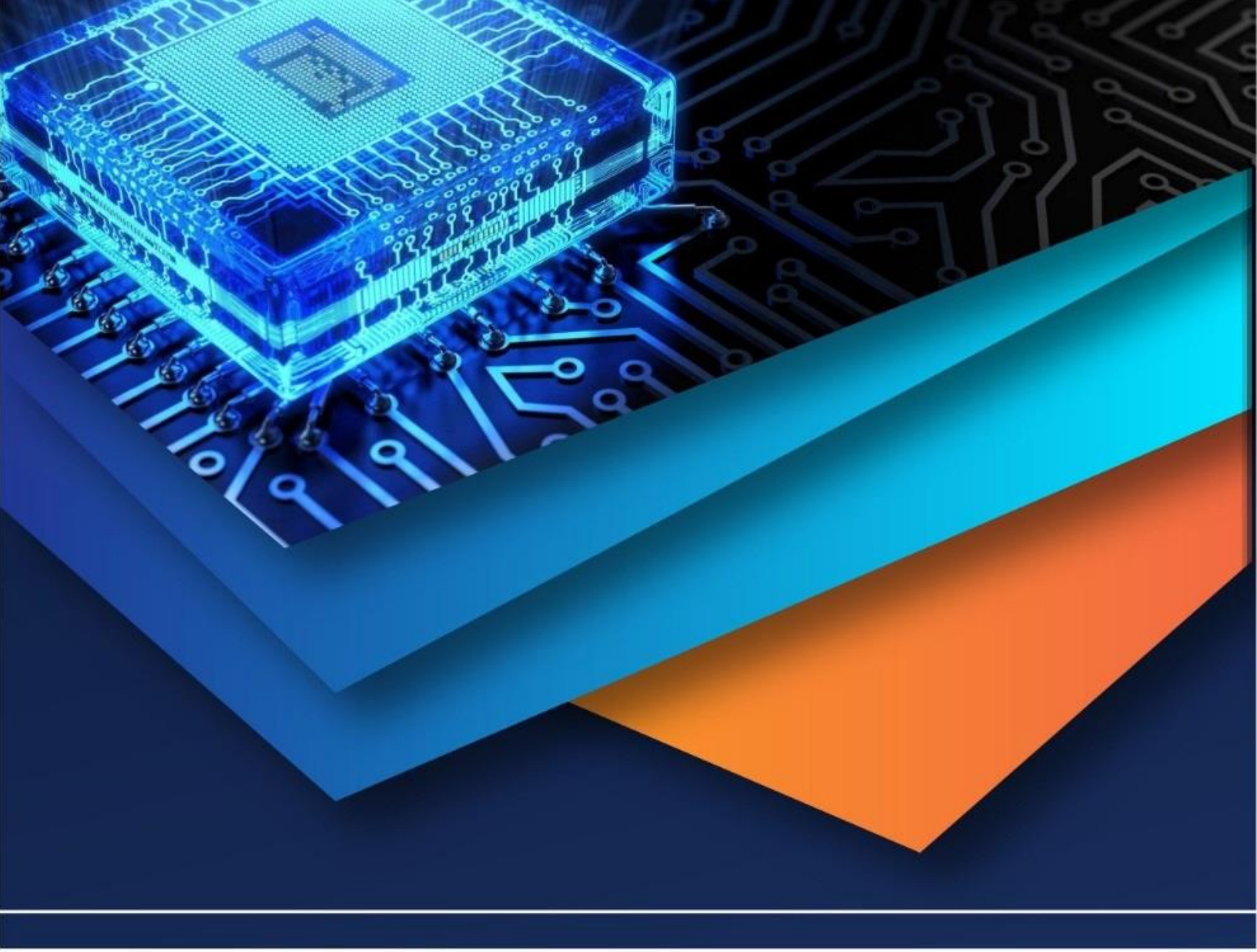
The Fig 5 shows the snapshot of next day weather prediction and also it provides weather predictions for next 5 days by sliding the screen.

VII. CONCLUSION

The physical models of the atmosphere can be accurately solved, but the instability of the physical models causes errors to accumulate quickly. On the other hand, machine learning algorithms are robust to perturbations in initial conditions, and over longer periods of time, perhaps our models would outperform the professional weather forecasting services. The obtained results are based on the attribute selection process on the datasets obtained from Google API that has different number of attributes. Machine Learning technique gives one of the best ways for weather prediction using previous datasets to analyze the future temperature, humidity, wind speed and pressure. By using decision tree of ID3 algorithm, this application predicts weather for next 5 days.

REFERENCES

[1] Abramson, Bruce, et al."Hailfinder: A Bayesian system for forecasting severe weather."International Journal of Forecasting12.1 (1996): 57-71.
[2] Cofno, Antonio S., et al. "Bayesian networks for probabilistic weather prediction."15th Eureopean Conference on Artificial Intelligence (ECAI). 2002.
[3] Krasnopolsky, Vladimir M., and Michael S. FoxRabinovitz. "Complex hybrid models combining deterministic and machine learning components for numerical climate modeling and weather prediction."Neural Networks19.2 (2006): 122-134.
[4] Lai, Loi Lei, et al. "Intelligent weather forecast."Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on. Vol. 7. IEEE, 2004.
[5] Radhika, Y., and M. Shashi. "Atmospheric temperature prediction using support vector machines."International Journal of Computer Theory and Engineering1.1 (2009): 55.
[6] Beniwal, S. &Arora, J.—Classification and feature selection techniques in data mining, International Journal of Engineering Research & Technology (IJERT), 2012.
[7] Sarah N. Kohail, Alaa M. El-Halees,—Implementation of Data Mining Techniques for Meteorological Data Analysis, IJICT Journal Volume 1 No. 3, 2011.
[8] S. Kannan , SubimalGhosh,—Prediction of daily rainfall state in a river basin using statistical downscaling from GCM output, Springer-Verlag, July- 2010.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY