2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI) | 979-8-3503-3384-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICDSAAI55433.2022.10028870

Weather Based Future Rain Prediction Using Machine Learning with Flask Framework

R.Sathya, CSE
SRM Institute of Science and Technology Ramapuram Campus, Chennai
sathyar@srmist.edu.in

Arpit Rastogi. CSE
SRM Institute of Science and Technology Ramapuram Campus, Chennai
ar4650@srmist.edu.in

Ankit Kumar, CSE
SRM Institute of Science and Technology Ramapuram Campus, Chennai
ak4908@srmist.edu.in

Shubham Singh, CSE
SRM Institute of Science and Technology Ramapuram Campus, Chennai
ss8233@srmist.edu.in

***Abstract*— Changes in weather patterns have an impact on a wide range of industries. Weather prediction that is intelligent and sophisticated is critical for reducing the influence of weather patterns. Agriculture is a big industry that is impacted by changing weather patterns. The amount of sunlight and rain that falls on particular crops has an impact on their production. To anticipate crop yield, various machine learning approaches have been developed. The major goal of this project is to create a smart weather forecast system that will improve crop output and raise agricultural profitability by utilizing various machine learning techniques. Datasets are gathered from a variety of sources and analyzed from a variety of angles in order to identify the many weather patterns that affect the crop yield. Currently, a variety of agricultural chores are carried out manually. Weather patterns are anticipated with the help of satellites orbiting the globe. The information is acquired from the satellites and analyzed by the competent authorities in the meteorological departments, who then take the necessary actions. All of this necessitates a significant amount of human effort and time. weather prediction is complete.**

Keywords—Datasets* sophisticated, *Machine Learning, Flask, Weather, Rain Prediction

I. INTRODUCTION

According to an IMD report, accurate weather forecasting benefits not just our daily life, but also food security and disaster management. The good news for India's monsoon-dependent economy is that we are becoming more adapt at forecasting. New technologies like the Internet of Things (IoT) and Artificial Intelligence (AI) are assisting meteorologists in better predicting agricultural output and natural disasters. In early 2020, Google researchers published a report claiming to be able to predict weather six hours ahead of time. Weather is an important component of a person's life since it can affect their health. Weather forecasting is meteorologists attempting to forecast weather conditions in the future as well as weather circumstances that may be expected. Temperature, pressure, humidity, dew point, rainfall and dataset size are used to calculate climatic condition.

Temperature, pressure, humidity, dewpoint, precipitation, and rainfall are solely taken into account for experimental purposes. Weather forecasting is the application of science and technology to anticipate the state of the atmosphere at any given time period. Weather forecasting can be done in a variety of ways. The importance of weather forecast alerts is that they can be used to prevent the loss of lives and the destruction of the environment. Pattern recognition was commonly utilized in historical weather forecasting methods, which meant that they relied on observing patterns of events. For example, if the previous day's sunset was exceptionally red, the following day's weather was deemed to be pleasant. All of the predictions, however, turn out to be incorrect.

Weather forecasting is essentially the art of predicting the weather in the future. Weather using historical data such as temperature, humidity, dew, wind speed and direction, precipitation, haze and air content, solar and terrestrial radiation, and so on. The weather forecast has a significant impact on people's lives. After the data has been collected, it must be trained. To begin, the data must be trained. We used 15-20%

979-8-3503-3384-8/22/$31.00 © 2022 IEEE

of Weather forecasting is the use of science and technology to predict the state of the atmosphere at any Regression Algorithm and the Nave Bayesian This prediction was made using a classification algorithm. This project employed Python, NumPy, Jupiter Notebook, Spyder, and Panda. This work is broken down into three: one for gathering, inspecting, and cleaning weather data, another for refining features and fitting the data to a Linear Regression and Nave Bayesian model.

II. LITERATURE SURVEY

Many research articles have been published for the prediction of weather using machine learning with flask. The study in [1] provides a new strategy for developing a service-oriented architecture for weather information systems that use data mining techniques to forecast weather. Algorithms for Artificial Neural Networks [12] [13] and Decision Trees [14] [16], as well as meteorological data obtained at a specific moment, can be used to do this. The best results for generating categorization rules for the mean weather variables came from the algorithm. Analysis on Weather Forecasting and Techniques [11]' concluded that artificial neural networks and the notion of fuzzy logic provide the best answer and prediction.

The major research [2] looked into the problems with weather prediction. The most basic weather forecast isn't always accurate. The next day's temperature forecast is usually within two degrees of the actual temperature. This accuracy isn't bad, though, because forecasts are produced for a longer length of time. Temperature projections, for example, are less accurate in a location like New England, where temperatures vary widely, than in the tropics. Another study, called "Current weather prediction," [3] used numerical methods to estimate what is most likely to occur based on current atmospheric circumstances. For instance, if a forecaster reviews three numerical models and two of them indicate that a storm would hit a given location, the forecaster will most likely predict that the storm will hit that location.

Bogdan Bochenek , and Zbigniew Ustrnul [4] Stated that it was also possible to extract the most commonly examined meteorological fields (wind, precipitation, temperature, pressure, and radiation), methods Preetham.M.S, Rohan Venkatesh Nayak, Uday.B, Radhika.T.V [5], used a Model of a Recurrent Neutral Network Using Machine Learning Techniques for Intelligent Weather Forecasting . The main advantages of this work are that is low cost and less power consumption increased privacy concerns. Vamshi K, Sachin Kumar S [6], employed the advantages of reent algorithms to predict real time rain locally. The disadvantages are season wish rain prediction. Urmay Shah ,Sanjay Garg ,Neha Sisodiya , Shashikant Sharma [7] Used Machine Learning and Forecasting Techniques to Improve Rainfall Prediction Accuracy. Predictions are generated by physical models based on a series of equations that anticipate future Rainfall in the dynamical approach, when it comes to long-term forecasting, it falls short. Naveen L,Mohan HS [8],using Big Data Analytics, analysed the impact of weather forecasting through deep learning on agricultural crop model prediction .a Deep Neural Network Prediction [17] [18]

To increase national food security, policymakers rely on accurate estimates to make timely import and export decisions. The concern availability, timeliness, and quality of observational data; time constraints on forecast preparation; the nature and reliability of communication systems available for forecast dissemination. Current weather prediction employed numerical approaches to anticipate what is most likely to happen based on current atmospheric conditions. For example, if a forecaster examines three numerical models and two of them indicate that a storm will hit a specific location, the forecaster will most likely predict that the storm will hit that location. These numerical models perform well and are constantly.

III. EXISTING SYSTEM

Currently, various activities that are happening in the agricultural sectors are carried out manually. The weather patterns are predicted with the help of satellites revolving around the world. From the satellites, the information is collected and they are analyzed by the concerned authorities in the meteorological departments and necessary actions are taken. All this requires a lot of human effort and time and the results are not very accurate. Weather is a complex atmospheric process and it is going to change dynamically under various situations. So, accurate weather prediction is a complex task and the results can’t be predicted with 100% accuracy.

The use of science and technology to predict the state of the atmosphere at any particular time is known as weather forecasting [10]. Forecasting the weather can be done in a variety of ways. The significance of weather forecast notifications is that they can be used to avert deaths and property damage. In the past, weather forecasting methods focused on observing patterns of events, which meant that they relied on observing patterns of happenings.

For example, if the sunset the day before was especially red, the weather the next day was judged nice. All of the predictions, on the other hand, prove to be inaccurate. The practice of anticipating future weather conditions using past data such as temperature, humidity, dew, wind speed and direction, precipitation,

haze and air content, solar and terrestrial radiation, and so on is known as weather forecasting shown in fig 1. Weather forecasts have a big influence on people's lives. It is necessary to train the data after it has been acquired. This study is around the Linear Regression approach, which is utilized to forecast weather using these data.

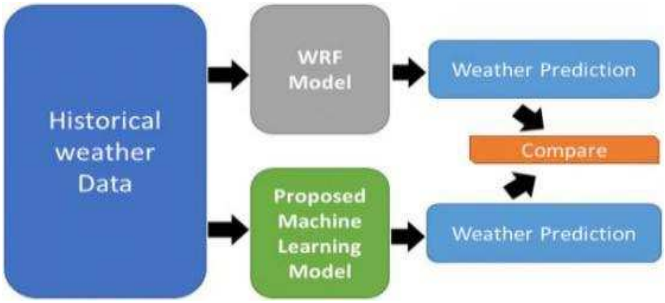

Fig. 1. Representation of existing weather prediction

Python is a high-level, interpreted programming language. Its formatting is simple and straightforward, and it typically uses English phrases instead of foreign language punctuation. It has a vast data mining and forecasting library. Jupiter Notebook/ Spider/ Pycharm: Jupiter Notebook/ Spider/ Pycharm: Jupiter Notebook/ Spider/ Pycharm is a cross-platform integrated development environment (IDE) for scientific Python programming that is free and open source. Spyder is compatible with a variety of well-known applications as well as open source software. Numpy: The front-end of the system was built with Numpy. Pandas was used to perform data preprocessing and statistical analysis. Matplotlib. For the graphical depiction of our forecast, Matplotlib was employed.

IV PROPOSED SYSTEM

The current system is developed to automate the process happening in the weather prediction and improve the accuracy of the overall system. Datasets are collected from various sources, pre-processing techniques are applied in order to remove any null values and various machine learning techniques are employed in order to predict the weather conditions under various circumstances and make the correct prediction. Supervised Classification algorithms like SVM and Decision tree algorithms are employed in order to predict the weather conditions. Essential features are extracted [11] and the machine learning models [12] are trained in order to make the correct prediction. The whole process is going to reduce the overhead on the farmers and it is going to increase the overall revenue in the agricultural sector.

The five models used to estimate rainfall levels in five, ten, and fifteen minutes were Random Forest, Back-propagation Neural Networks, Support Vector Machines, and Networks [15] are all examples of auto-regressive and moving average models. Weather data from the Wuhan region of China was used to test the forecast models in 2015 and 2016. Wind speed, wind direction, temperature, humidity, pressure, amount of rainfall, and radiation were among the features concentrated in the dataset. The LSTM-Networks-based forecast model yielded positive results. On the other hand, Aswin, Geetha, and Vinaya kumar [9] suggested a method for predicting monthly rainfall based on an model and a ConvNet model. Microwave data, infrared data, and rain gauge observations were used to extract precipitation estimation features. Weather data from the Global Precipitation Climatology Project was used to train and test the models. The RMSE and Mean were found to be significant. . The results showed that the RMSE and Mean Networks models yielded similar results. Networks are tasked with prediction monthly rainfall. India from was used as the climatic dataset.

A. Architecture Diagram

Fig 2 depicts an architecture diagram used to extract the whole and Fig 3 shows the flow diagram of the proposed work.

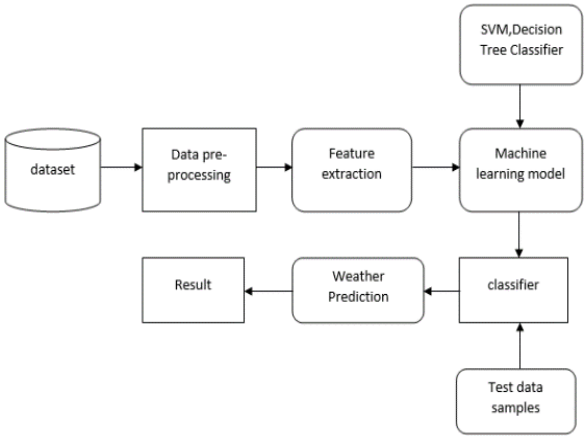

Fig 2. To show the flow architecture diagram

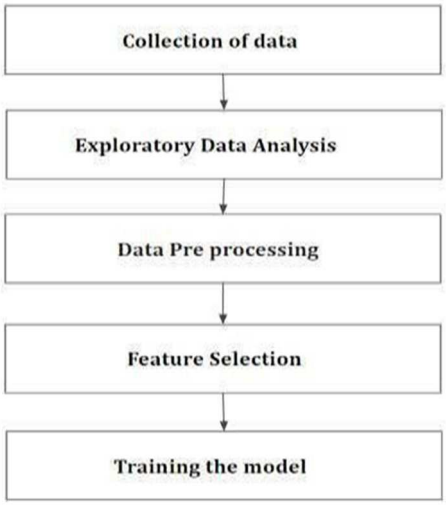

Fig 3 Data Flow Diagram

B. Description

Website*:* As shown in fig. 2*,* the web page to check the AccuWeather collects the most up-to-date and comprehensive weather data in order to provide forecasts with Superior Accuracy. Forecasts are precise for every area on the planet and go out further than any other source..

Dataset*:* As data set is a grouping of data. A data set corresponds to one or more database tables in the case of tabular data, with each column of a table representing a single variable and each row representing a single row, representing a specific record of the data set in question.

System Model: As shown in fig 2, the system consists of input datasets of proven and unproven remedies, analysis of the data as proven/unproven by training and testing using Logistic Regression and Decision Tree, and Naive Bayes going to detect whether the remedy is unproven or not, and final output.

Process: The fig *2* is to show system with architecture Diagram show the dataset with preprocessing Feature extraction and Machine learning model is classifier. Finally*,* the data which we process in training and testing using two machine earning models to predict the weather of next day using some information to get binary output as sunny day or rainy day. It is shown as a result of weather prediction using html language with values such as date, location, wind gust direction, minimum temperature, maximum temperature, Rainfall, Sunshine, Humidity and Cloud.

V MODULES USED

A. Data collections

A database collection is a collection of computer-controlled data items that may be studied and forecasted as a single unit in machine learning. The information we gather is intended for and understandable by a machine that does not interpret information in the same way that humans do. Using Logistic regression, Decision trees, and Navie Bayes to gather data, handle missing data, move data through feature extraction, identify essential features, and separate data from training and testing sets.

B. Selection of features

Feature selection is a technique for decreasing the number of input variables in a predictive model. We used this strategy to try to reduce the number of input variables in order to reduce the computing costs of modeling and in certain cases, to improve the model's performance mathematical. Using mathematics to assess the link between each input variable and target variations and selecting the input variables and target variables with the strongest correlations are examples of options-based selection procedures.

C. Model training

Model training in machine learning is a strategy for feeding information to the ML algorithm that enhances identification and helps us learn good-supervised reading. A database training model was used to train the ML algorithm. Specimen data is transported with corresponding sets of I/O data that influence the outcome. The training replica is used f to consolidate the analyzed result by removing the sample. As a result of the merger, the model has changed. Data sets are used to train the machine learning algorithm during this time. This is where learning takes place. With compatible training, the ML replica's level of prediction can be increased. The mass of the model should be begun at random.

D. Evaluating the model used

Model testing will be used to assess the model's accuracy on future data (invisible or outside sample). The methods for measuring model performance are divided into two groups: date, location, minimum temperature, maximum temperature, humidity, and cloud. Both methods evaluate model performance using a test set (data that isn't from the model).

E. Training Model

The final machine learning model will be used to anticipate when consumers will be provided new information. That is, we require a model that can predict the projected outcome using new instances as input data. With the help of Trained-Model, we will be able to recognize rain predictions and show them on websites. Data preparation, algorithm steps and algorithm correction are all part of the procedure. We gather data at work, put in the hours, and then find out

how to prepare the data, which method to use, and how to present it.

VI IMPLEMENTATION

We predict verified and unconfirmed data sets that use Logistic regression, decision tree algorithm as in fig 4

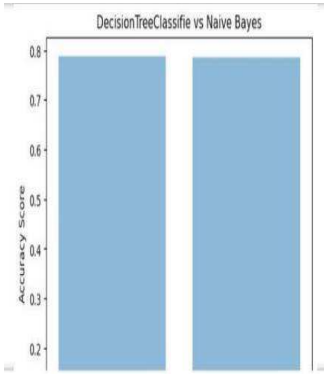

Fig. 4. Comparison in two algorithms

A. Data Input

Every piece of data contains a binary string with a length of twenty. The representation technique will be similar to that of a This list of python to show binaries represented as 1s and 0s . A wide range of such lists can be found in the trial information for training. [batch_ dimension, order_ length, information_ dimension] are all required for logistic regression (variable). The batch_ dimension will be addressed later in this circumstance. However, because we have a data collection with rain prediction data, order_ length has been set to 100.

B. Operation

Fig 5 depicts the webpage for weather forecasting. In most circumstances, having values for temperature at various times, humidity at various times, cloud at various times, date, time, and whether or not it is raining. This embedding is unique. We created a single equally hot thermal representation throughout the training installation cycle. Connected width backend with some Hypertext markup language pages having input and display, the final result, is raining or sunny day. Train and test the model and give backend call for it.

.

Fig.5. predicting the weather

VII WORKING PROCESS

A. Project Execution Procedures

First, the necessary packages, such as numpy, pandas, matplotlib, seaborn, nltk were installed and train test split, classification report, and accuracy score were measured. To read – pd.read_csv ("weatherAUS .csv"). We present a graph shown in Fig. 6 outlines the accuracy of the two models that will be used to verify the rain prediction in the project. Weather based rain prediction datasets. in Fig 7, will aid in the display of text data, with the size of each word indicating its frequency or relevance. The comparison of the two models employed as decision tree classification technique is shown in Fig 8. Any decision can be made at a decision node, which has numerous branches. The features of the given dataset are used to make judgments or run tests. It's a diagram that shows how to get all of the information you need. The Nave Bayes Classifier is a simple and effective classification method for developing fast machine learning models that can make quick predictions. In Nave The Bayes Classifier is a simple and successful classification method for developing fast machine learning models that can make quick predictions. It's a probabilistic classifier, meaning it makes predictions based on the probability of an object. The Nave Bayes Algorithm is commonly used for spam filtration, sentiment analysis, and article classification. Compare the algorithm and get the best result with most accuracy. Fig 9 shows heatmap of the model

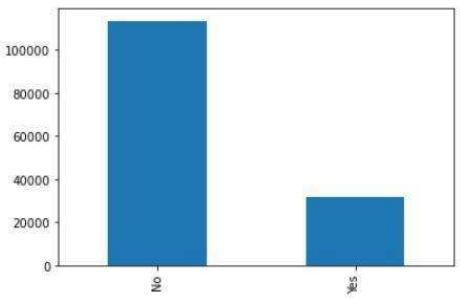

Fig. 6. Accuracy of rain predication

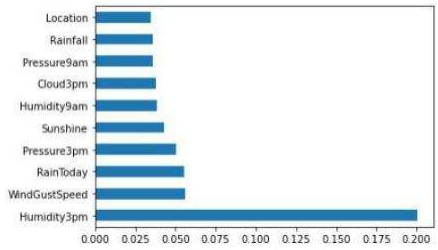

Fig. 7.Implementation of used model

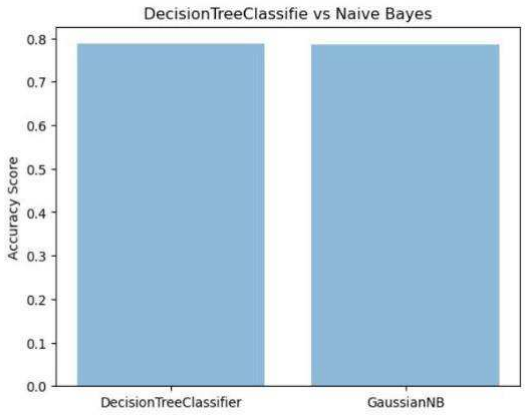

Fig. 8 accuracy of models

Fig 9 Heatmap

VIII RESULT AND DISCUSSION

To guess the given solution for the Proven or Unproven through this system, a smart weather prediction system has been developed using various algorithms are trained in such a manner that it is going to improve the overall accuracy of the system and improve the revenue in the agricultural sector . The weather patterns are analyzed under various circumstances and machine learning models are trained in order to make the correct predictions.

The logistic regression model has an accuracy score of 0.84. In terms of predictions, the model performs admirably. The model shows no evidence of underfitting or overfitting. The accuracy score of cross-mean validation is roughly identical to the accuracy score of the original model. Cross-validation may not improve the model's accuracy in India as a result. We used data from the Karnataka State Natural Disaster Monitoring Centre, district-level Monthly Actual Average Rainfall, temperature, and relative humidity data, as well as agriculture data from Karnataka's five districts: Bangalore, Belgaum, and Bellary, in this work.

IX CONCLUSION AND FUTURE WORK

Thus, through this system, a smart weather prediction system has been developed using evidence of underfitting or overfitting. This demonstrates that the model can generalize to new data. The accuracy score of cross-mean validation is roughly identical to the accuracy score of the original model. Cross-validation may not improve the various algorithms for machine learning Machine learning algorithms are trained in such a way that they improve the system's overall accuracy and performance. The revenue in the agricultural sector. The weather patterns are analyzed under various circumstances and machine learning models are trained in order to make the correct predictions. The logistic regression model has performs admirably. an accuracy score of 0.84. In terms of predictions, the model shows no model's accuracy as a result.

Future Work

The future work of this project is that it can be improved by making the website more attractive and connect it with the maps application such as google maps API, OLA, UBER which also help it to guide about the weather in coming areas and using some other models the things to accuracy predicted. This is the Solution of farmer to get the result to help in which crop to be applied and yield quality can b e improved and The logistic regression model has an accuracy score of 0.84. In terms of predictions, the model performs admirably. The model shows no evidence of underfitting or overfitting. This demonstrates that the model can new data. The accuracy score of cross-mean validation is roughly identical to the accuracy score of the original model.

REFERENCES

[1] *Pandey, Swati & Sharma, Shruti & Kumar, Shubham & Bhatt, Kanchan & Arora, Rakesh. (2021). Analysis of Weather Forecasting Techniques. International Journal of Scientific Research. 7. 80-85. 10.32628/CSEIT217421.*

[2] Biswas, Munmun & Dhoom, Tanni & Barua, Sayantanu. (2018). Weather Forecast Prediction: An Integrated Approach for Analyzing and Measuring Weather Data. International Journal of

Computer Applications. 182. 20-24. 10.5120/ijca2018918265.

[3] Jana Sillmann, Thordis Thorarinsdottir, Noel Keenlyside, Nathalie Schaller, Lisa V. Alexander, Gabriele Hegerl, Sonia I. Seneviratne, Robert Vautard, Xuebin Zhang, Francis W. Zwiers, Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities,Weather and Climate Extremes, Volume 18,2017,Pages 65-74,ISSN 2212-0947, https://doi.org/10.1016/j.wace.2017.10.003

[4] Bochenek, Bogdan, and Zbigniew Ustrnul. 2022. "Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives" Atmosphere 13, no. 2: 180. https://doi.org/10.3390/atmos13020180

[5] Preetham M S , Prathap N Kashyap , Rohan Venkatesh Nayak , Uday.b , Radhika.t.v, Intelligent Weather Forecasting using Machine Learning Techniques, International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 03 | Mar 2020

[6] Vamshi K1, Sachin Kumar S2, Muralidhar B R3, Manjunath N4, Mrs. Savitha P5 , A Review on Rainfall Prediction using Machine Learning and Neural Network, International Research Journal of Engineering and Technology (IRJET) , Volume: 08 Issue: 03 | Mar 2021

[7] Shah, Urmay & Garg, Sanjay & Sisodiya, Neha & Dube, Nitant & Sharma, Shashikant. (2018). Rainfall Prediction: Accuracy Enhancement Using Machine Learning and Forecasting Techniques. 776-782. 10.1109/PDGC.2018.8745763.

[8] Naveen L1 , Mohan H S2 , Analyzing Impact of Weather Forecasting Through Deep Learning in Agricultural Crop Model Predictions, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 23 (2019) pp. 4379-4386

[9] Aswin, S. et al. "Deep Learning Models for the Prediction of Rainfall." 2018 International Conference on Communication and Signal Processing (ICCSP) (2018): 0657-0661.

[10] Arne Sundae, Predict Weather, 2015. Scott C. James is a writer. "A machine learning framework to forecast wave conditions," Yushan Zhang and Fearghal O'Donncha, Coastal EngineeringVolume 137, July 2018, Pages

[11] Sathya, R., Rawat, D., Mondal, A., Choudhary, S., & Jain, A. (2019). Economically efficient data feature selection using big data analysis. International Journal of Innovative Technology and Exploring Engineering, 8(7), 983–987.

[12] Md Mosharaf Hossain, A H M Jakaria, and Mohammad Ashiqur Rahma 2018. A Case Study in Tennessee Using Machine Learning for Smart Weather Forecasting. ACM Mid-Southeast Conference Proceedings (Mid-Southeast'18). New York, NY, USA: ACM.

[13] Sathya, R., Saleena, B. CNN-MAO: Convolutional Neural Network-based Modified Aquilla Optimization Algorithm for Pothole Identification from Thermal Images. SIViP (2022). https://doi.org/10.1007/s11760-022-02189-0

[14] R. Sathya, Muppala Nikhil, K.S.S. Shanmukha Srinath, Mula Ranpal Reddy Real Time Image Classification and Detection of Potholes using Adam's Optimizer with CNN Framework, International Journal of Early Childhood Special Education (INT-JECSE) DOI: 10.9756/INT-JECSE/V14I2.400 ISSN:1308-5581 Vol 14, Issue 02, 2022

[15] YomgviChen and WengianZhang, Mathematical Models of Multifactororial Decision and Weather Forecast," IFAC Proceedings, Volume 16, Issue 13, July 1983, Pages 265-269J.

[16] Sathya, R., Saleena, B. A Framework for Designing Unsupervised Pothole Detection by Integrating Feature Extraction Using Deep Recurrent Neural Network. *Wireless Pers Commun* (2022). https://doi.org/10.1007/s11277-022-09790-z

[17] Mr. Sunil Navadia, Mr. Jobin Thomas, Mr. Pintukumar Yadav, Ms. Shakila Shaikh, "Weather Prediction: A novel approach for measuring and analyzing weather data", International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), (I-SMAC 2017), IEEE, pp 414-417

[18] Mehrnoosh Torabi, Sattar Hashemi, "A Data Mining Paradigm to Forecast Weather", The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012),IEEE, pp 579-584