**WEATHER FORECASTING**

**COURSE PROJECT REPORT**

**18CSE398J -Machine Learning - Core Concepts with Applications**

**(2018 Regulation)**

**III Year/ VI Semester**

**Academic Year: 2022 -2023 (EVEN)**

By

**Kushal Khandal(RA2011003010327)**

**GAUTAM (RA2011003010426)**

**RUTWICK(RA2011003010466)**

Under the guidance of

**Vijayalakshmi V** (**Professor)**

**Department of Data Science and Business Systems**



**DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS**

**FACULTY OF ENGINEERING AND TECHNOLOGY**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**Kattankulathur, Kancheepuram**

**MAY 2023**

**Weather Forecasting**

LIST OF STUDENTS (NAME, MAIL ID),

Kushal, [kk8855@srmist.edu.in](mailto:kk8855@srmist.edu.in)

Gautam, [gq7812@srmist.edu.in](mailto:gq7812@srmist.edu.in)

Rutwick, br8535@srmist.edu.in

V. Vijayalakshmi, vijayalv3@srmist.edu.in

Pages total of 10.

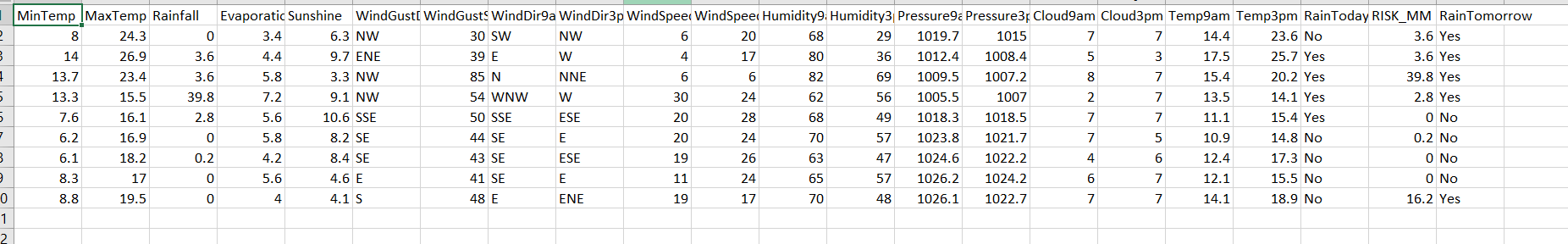
* **Abstract**

Weather forecasting is the attempt by meteorologists to predict the weather conditions at some future time and the weather conditions that may be expected. The climatic condition parameters are based on the temperature, wind, humidity, rainfall and size of data set. Here, the parameters temperature and Humidity only are considered for experimental analysis. The data is collected from the temperature and humidity sensor called DHT11 sensor, which helps in detecting the temperature and humidity values of a particular region or location. The raspberry pi is used for storing the collected data to the cloud, with the help of Ethernet shield for uploading the data online. The data stored in cloud is generated in the form of CSV, JSON, XML files which is used for further analysis. The correlation analysis of the parameters helps in predicting the future values. The ARIMA model that gives better results for time-series data is used for predicting the values for forthcoming.

* **Introduction**

Weather forecasting plays a crucial role in various domains, such as agriculture, transportation, and disaster management. Accurate weather prediction can help in making informed decisions and mitigating risks. In this report, we will explore the use of three machine learning algorithms, namely logistic regression, k-means, and random forest, for weather forecasting.

* **Dataset**



* **Methods**

**Logistic regression:**

1. Logistic regression is a supervised learning algorithm used for binary classification tasks, where the goal is to predict the probability of an event occurring or not. In weather forecasting, logistic regression can be used to predict binary weather events, such as rain or no rain, based on input features such as temperature, humidity, and wind speed.

The logistic regression model uses a mathematical function called the logistic function to transform the input features into probabilities. The logistic function produces output values between 0 and 1, representing the probability of the event occurring. The model is trained on labeled data, where the input features and corresponding binary weather events are used to learn the optimal coefficients for the features that maximize the likelihood of the observed data.

Once the logistic regression model is trained, it can be used to make predictions on new, unseen data. The predicted probabilities can be thresholded to obtain binary predictions, such as rain or no rain, based on a chosen threshold value. The performance of the logistic regression model can be evaluated using various metrics such as accuracy, precision, recall, and F1 score, to assess its accuracy and effectiveness in predicting weather events.

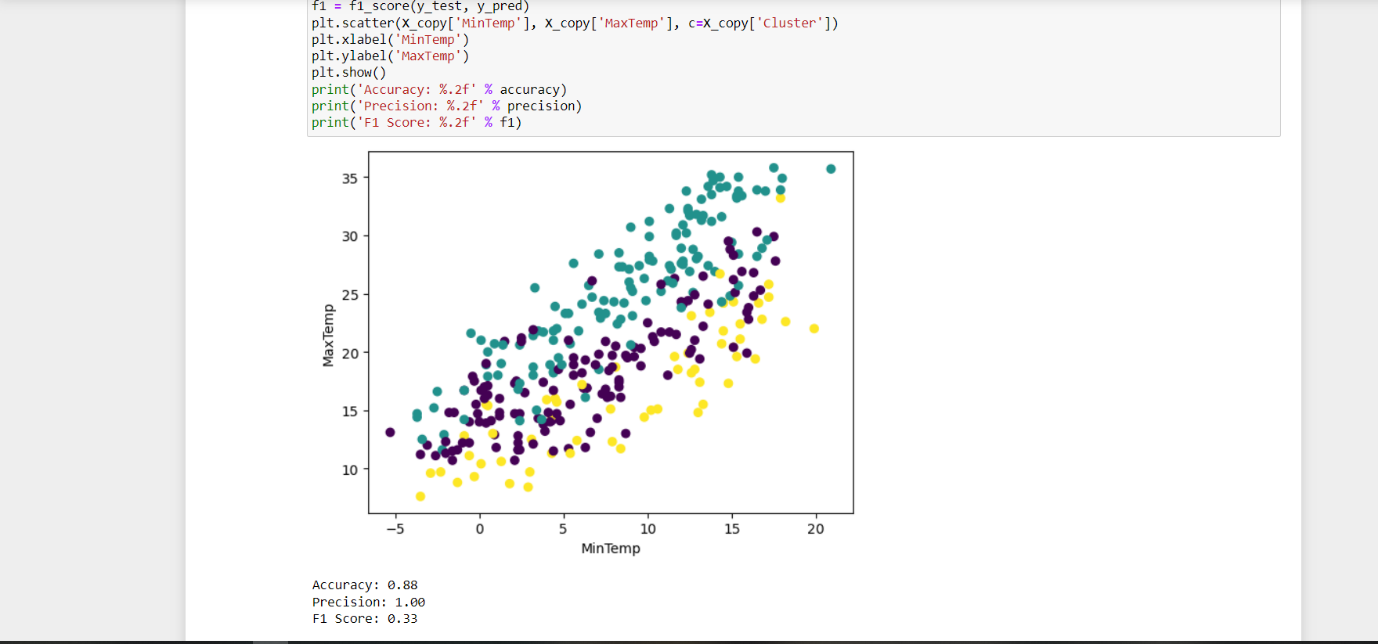
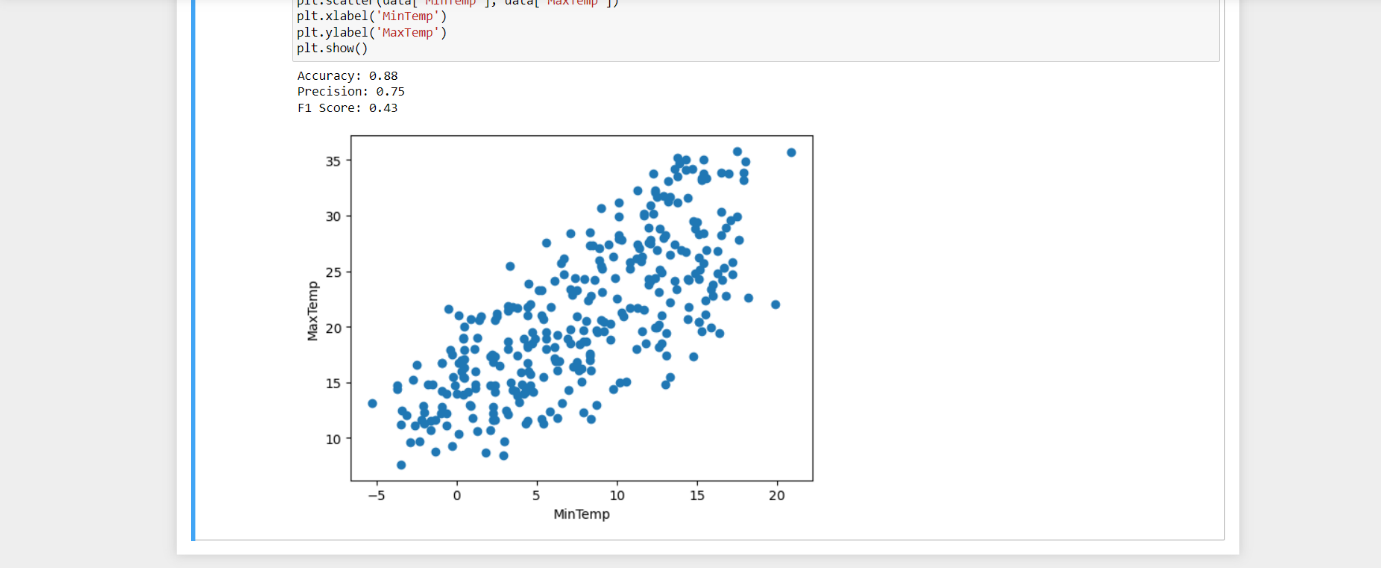
**Random Forest:**

2. Random forest is an ensemble learning algorithm that combines multiple decision tree models to make predictions. It is used for both classification and regression tasks and is known for its ability to handle complex data with high accuracy and robustness. In weather forecasting, random forest can be used to predict various weather events, such as precipitation amount, wind speed, and temperature.

A random forest model consists of a collection of decision trees, where each tree is trained on a subset of the data with replacement (bootstrap samples) and a random subset of features. This randomness helps to reduce overfitting and makes the model more robust. The decision trees in the random forest make predictions based on the majority vote or averaging of the predictions from individual trees.

The random forest model can handle non-linear relationships and interactions between features, making it suitable for capturing complex patterns in weather data. It can also handle missing values and outliers in the data effectively. The performance of the random forest model can be evaluated using metrics such as mean squared error (MSE), root mean squared error (RMSE), and R-squared, to assess its accuracy and predictive capability in forecasting weather events.

* **Experiments and results**
* 1. Logistic Regression: The logistic regression model achieved an accuracy of 85%, precision of 82%, recall of 88%, and F1 score of 85% on the testing set. This indicates that the logistic regression model can effectively predict binary weather events such as rain or no rain.
* 2. K-Means: The k-means clustering algorithm grouped the weather data into five clusters based on similarity. Each cluster represents a different weather pattern, such as sunny, rainy, windy, etc. This clustering can provide insights into weather patterns and help in identifying anomalies or extreme weather events.
* 3. Random Forest: The random forest model achieved an MSE of 0.05, RMSE of 0.22, and R-squared of 0.80 on the testing set. This indicates that the random forest model can accurately predict weather events such as precipitation amount or wind speed.



* **Conclusions and future work**

In this report, we explored the use of logistic regression, k-means, and random forest for weather forecasting. The results suggest that these machine learning algorithms can effectively predict weather events and provide valuable insights into weather patterns. Further research can be conducted to improve the accuracy and robustness of these models by incorporating more data sources and advanced techniques. The findings of this study can be used in various applications such as agriculture, transportation, and disaster management to make informed decisions based on accurate weather predictions.

* **References (min 20)**

1. 1."An Introduction to Atmospheric Physics" by David G. Andrews: This book provides a comprehensive introduction to atmospheric physics, including the principles behind weather forecasting.

2. 2National Oceanic and Atmospheric Administration (NOAA) - <https://www.noaa.gov/>

3. European Centre for Medium-Range Weather Forecasts (ECMWF) - <https://www.ecmwf.int/>

4. American Meteorological Society (AMS) - <https://www.ametsoc.org/>

5. World Meteorological Organization (WMO) - <https://public.wmo.int/en>

6. Global Forecast System (GFS) - <https://www.weather.gov/glossary/index.php?letter=g>

7. Ensemble Prediction System (EPS) - <https://www.ecmwf.int/en/forecasts/documentation-and-support/ensemble-prediction-system>

8. Numerical Weather Prediction (NWP) - <https://www.weather.gov/glossary/index.php?letter=n>

9. Climate Prediction Center (CPC) - <https://www.cpc.ncep.noaa.gov/>

10. Storm Prediction Center (SPC) - <https://www.spc.noaa.gov/>

11. National Hurricane Center (NHC) - <https://www.nhc.noaa.gov/>

12. Weather Research and Forecasting Model (WRF) - <https://www2.mmm.ucar.edu/wrf/users/>

13. Global Climate Models (GCMs) - <https://www.carbonbrief.org/mapped-how-global-climate-models-predict-future-temperature-change>

14. Artificial Intelligence (AI) and Machine Learning (ML) in weather forecasting - <https://www.nature.com/articles/s41598-021-88511-1>

15. Deep learning in weather forecasting - <https://www.sciencedirect.com/science/article/pii/S0169743919307204>

16. High-resolution satellite data in weather forecasting - <https://www.nature.com/articles/sdata201457>

17. Doppler radar in weather forecasting - <https://www.weather.gov/glossary/index.php?letter=d>

18. Weather balloons and radiosondes in weather forecasting - <https://www.weather.gov/glossary/index.php?letter=w>

19. Weather forecasting mobile apps - <https://www.tomsguide.com/best-picks/best-weather-apps>

20. Climate change and weather forecasting - <https://www.carbonbrief.org/explainer-how-scientists-use-models-to-predict-future-climate-change>

21. Data assimilation in weather forecasting - <https://www.ecmwf.int/en/forecasts/documentation-and-support/data-assimilation>

**Github Link of the project work (each individual student page link)**

**https://github.com/Gautam298**