**Project Report**

**Title of the project** : Rainfall Prediction

**Name of the Student** : Hemanth Phani Sai. Mothukuri

**Regd. No** : AP18110010535

**Semester/Year**: VI / 3rd

**Project Category (UROP/ CAPSTONE/IDS)** : IDS

**1.Objective of the Project *:-***

I will show you how to compare different algorithms and choose the best ones. Instead of considering the single algorithm for the entire project and then realizing that the performance is not good, we first check the performance of some algorithms and then decide which one is useful for the project.my goal is to evaluate several classification algorithms and select the best one based on accuracy.

The aim of this project is to build and identify the best binary classification model predicting rainfall in Australia using two common algorithms (decision tree and random forest). The dataset used in this project describes variables such as temperature, wind and humidity with labels indicating whether it rained that day or the day after.

The accurate and exact predictions will help in developing the more appropriate strategies for agriculture and water reserves and will also be informed about the flood to implement precautionary measures.

the proposed study is too effective and efficient in predicting the rainfall with accuracy and precision.

**2.Introduction :-**

Machine learning is related to the study of the algorithms that enhance the efficiency of the machines/computers automatically through the training and testing of the machine/computers with certainly different variables. The machine learning is among the foremost favourable and fastest growing areas of technology. The computers work efficiently with different algorithms and functions. The machine learning is that the training the pc with certainly different algorithms to experience the machine in automatic smart processing. The machine learning enhances the efficiency and accuracy of the info processing and is employed during a wide selection of fields. The machine learning is developed with effectual algorithms that utilize a particular set of tools and functions to unravel the complex and large data. The machine learning is assisting within the diverse field; these are normally AI applications that are used for recognition and prediction like that of computer engineering and medical fields. The machine learning is popular among the fashionable technology and has many benefits. Machine learning develops some rules for the input file as discussed within the hybrid model that helps the machine to process the similar case each tie efficiently. It works on the prediction, and it's more important to understand that how variables the inputs into are moved into vectors. The machine learning has minimized the manual jobs for the folks that also could have the space for errors and inaccuracy

Classification is a process of categorizing a given set of data into classes, it can be performed on both structured or unstructured data. the process starts with predicting the class of given data points. The classes are often referred to as target, label or categories.

Classification Terminologies :-

* Classifier
* Classification Model
* Feature
* Binary Classification
* Multi-Class Classification
* Multi-Label Classification

Algorithm Selection :-

* Read the data
* Explore the data
* Split the data
* Train the model
* Classifier Evaluation
* Accuracy
* Precision
* recall
* F1 Score

Accuracy :-

It is the ratio of correctly predicted observation to the total observations.

Accuracy = TP + TN / TP + FP + FN + TN

Precision :-

precision is the fraction of relevant instances among the retrieved instances

Precision = TP / TP + FP

Recall :-

recall is the fraction of relevant instances that have been retrieved over the total number of instances

Recall = TP / TP + FN

F1 Score :-

The weighted average of the precision and recall.

F1 Score = 2 \* (Recall \* Precision) / (Recall + Precision)

**3. Literature Survey :-**

Rainfall prediction is not an easy job especially when expecting the accurate and precise digits for predicting the rain. The rainfall prediction is commonly used to protect the agriculture and production of seasonal fruits and vegetables and to sustain their production and quality in relation to the amount of rain required by them (Lima & Guedes, 2015).

The rainfall forecasting is prevailing as a popular research in the scientific areas in the modern world of technology and innovation; as it has a huge impact on just the human life but the economies and the living beings as a whole. Rainfall prediction with several Neural Networks has been analyzed previously and the researchers are still trying hard to achieve the more perfect and accurate results in the field of rainfall prediction (Biswas, et al., 2016).

Water as is one of the most useful resources of the earth. There is no human and living thing on earth that can survive without water. As, this precious resource is running out because of the increasing temperature of the earth and the unexpected and unappreciated climatic conditions due to global warming. (Mittal, Chowdhury, Roy, Bhatia, & Srivastav, 2012).

The rainfall prediction is also emphasized for its significance for the prediction of flood and consequently takes the precautionary measure to save the people from devastating destructions that a flood can cause (Hoai, Udo, & Mano, 2011). There are studies that outlined the significance of rainfall prediction in forecasting flood on the regions where there is heavy rain every year. The areas with high risk for flood are the vulnerable areas that need the rainfall forecasting not just to save a human life but to safe agriculture, water reservation and livestock (Fang & Zhongda, 2015).

**4. Proposed Method *:-***

Classification Algorithms Used :-

Decision Tree :-

The decision tree builds the classification model with in the tree structure. It utilizes the if – then rules which are equally exhaustive and mutually exclusive in classification.

Random Forest :-

Random decision trees or random forest are an ensemble learning method for classification etc. it operates by constructing a multitude of decision trees at training time and outputs the class that is the mode of the classes or classification of the individual trees.

***5.* Dataset *:-***

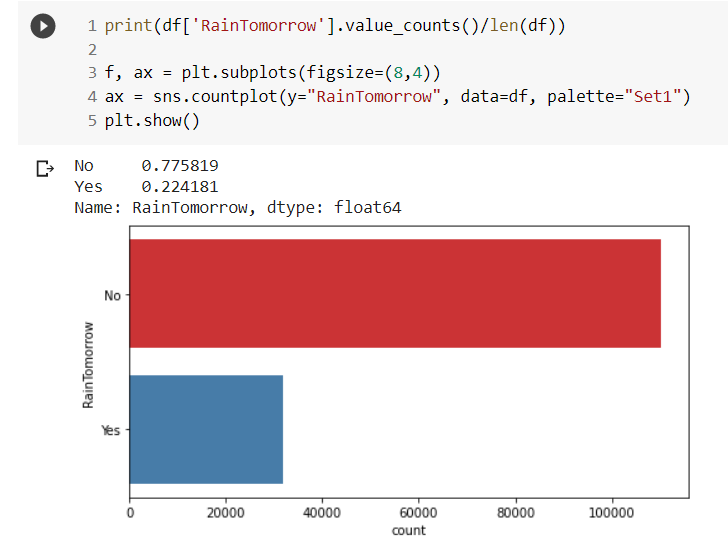
weatherAUS.csv

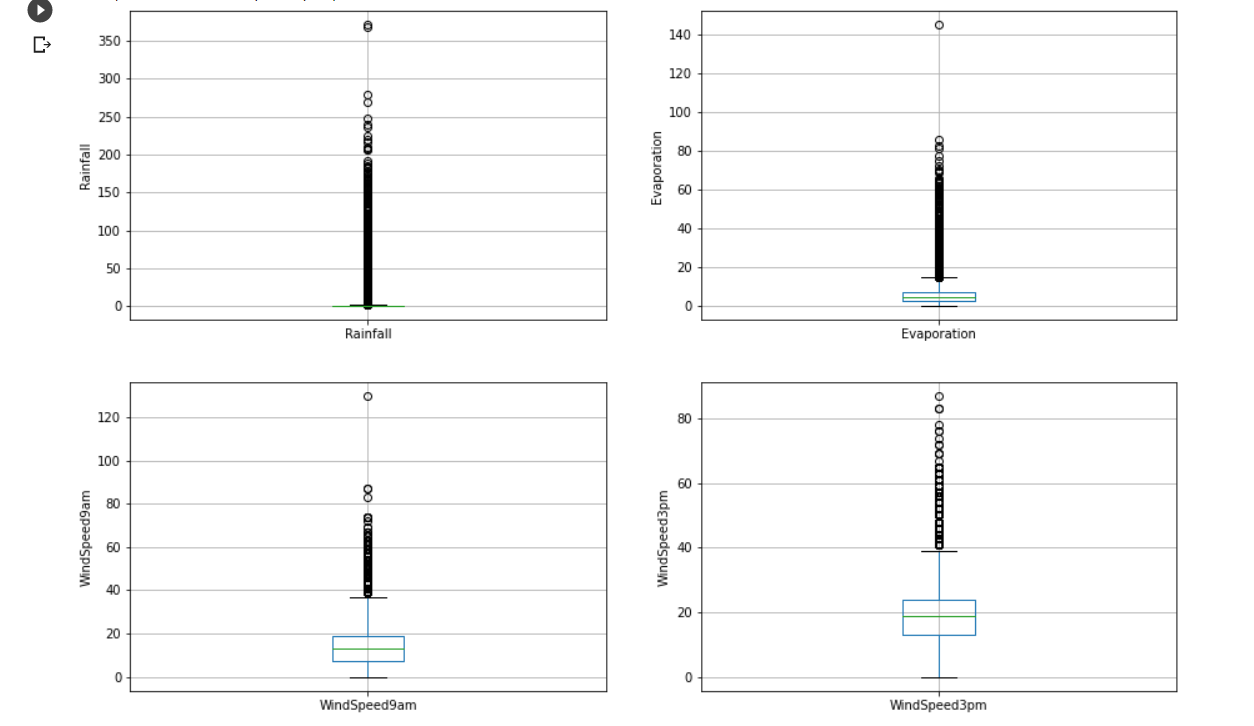
[**https://drive.google.com/file/d/1KfOo\_mAmq0swCaH9aPVovWeLguE1dLgH/view?usp=sharing**](https://drive.google.com/file/d/1KfOo_mAmq0swCaH9aPVovWeLguE1dLgH/view?usp=sharing)

**6. Results and Discussion *:-***

Climate change is usually a serious issue for whole world and making any prediction on it is now days pretty difficult and unpredictable. Climate change is because of the present warming trend is human expansion. Due to this air and oceans are warming, water level is rising and flooding and drought etc. One of the intense consequences thanks to this global climate change is on Rainfall. Rainfall prediction now days is an arduous task which is taking into the consideration of most of the main world-wide authorities.

In this paper core motive is to looking for the algorithm which provides us the great prediction of rainfall. Here are we took the rainfall data of India of past 23 year from the official site of Indian government. Below is that the table of the accuracy of the algorithms.





DECISION TREE

Model accuracy score: 0.7883

Training set score: 1.0000

Test set score: 0.7883

precision recall f1-score support

0 0.87 0.86 0.86 22098

1 0.52 0.54 0.53 6341

accuracy 0.79 28439

macro avg 0.70 0.70 0.70 28439

weighted avg 0.79 0.79 0.79 28439



RANDOM FOREST

Model accuracy score: 0.8186

Training set score: 0.8197

Test set score: 0.8186

precision recall f1-score support

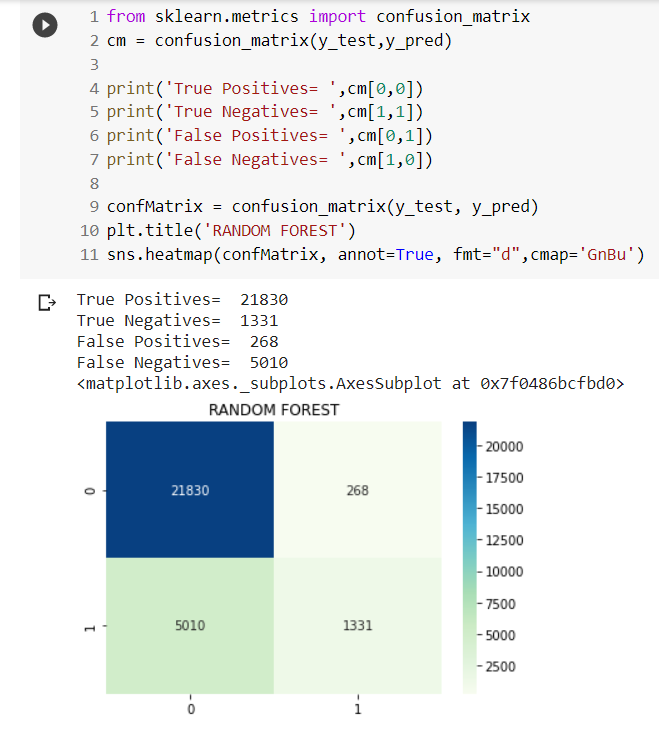
0 0.82 0.98 0.89 22098

1 0.82 0.24 0.37 6341

accuracy 0.82 28439

macro avg 0.82 0.61 0.63 28439

weighted avg 0.82 0.82 0.78 28439



**7. Conclusion :-**

We have used different machine learning algorithms to check the accuracy and precision of rainfall prediction. we've comparing Decision Tree, Random Forest classifiers. By seeing the result we have conclude that Random forest is the Machine learning algorithm which is suitable for rainfall prediction.

**8. References :-**

<https://medium.com/analytics-vidhya/classification-in-machine-learning-ed30753d9461>

<https://www.edureka.co/blog/classification-in-machine-learning/>

<https://morioh.com/p/8c87222937cb>

<https://www.tutorialspoint.com/fraud-detection-in-python>

<https://devincept.codes/ml/Classification%20Methods%20in%20ML/ml-classification.html>

<https://www.edureka.co/blog/classification-in-machine-learning/#:~:text=Neighbor%20Algorithm%20here-,Decision%20Tree,and%20mutually%20exclusive%20in%20classification.>

<https://ukdiss.com/examples/rainfall-prediction-machine-learning.php>

**Note:**

* **The same format needs to be used for report submission. (Word File only)**
* **Submit the plagiarism report as a separate copy. More than 15% plagiarism is not acceptable.**