JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

RAINFALL PREDICTION MODEL USING MACHINE LEARNING.

A JENNIFER MARY , GNANA SREE B

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

S.E.A COLLEGE OF ENGINEERING AND TECHNOLOGY,560049,INDIA.

ABSTRACT

Rainfall prediction plays a crucial role in raising awareness about the potential dangers associated with rain and enabling individuals to take measures for their safety. This study aims to utilize machine learning algorithms to accurately predict rainfall, considering the significant impact of scarcity or extreme rainfall on both rural and urban life. The complex nature of rainfall, influenced by various atmospheric, oceanic, and geographical factors, makes it a challenging phenomenon to forecast. This research employs data pre-processing techniques, outlier analysis, correlation analysis, feature selection, and several machine learning algorithms such as Naive Bayes (NB), Decision Tree, Support Vector Machine (SVM), Random Forest, and Logistic Regression. The study focuses on developing the most accurate rainfall prediction model by utilizing machine learning and feature selection techniques. The Artificial Neural Network (ANN) achieves a maximum accuracy of 90% and 91% before and after feature selection, respectively. Overall, this research demonstrates the effectiveness of different machine-learning techniques in predicting rainfall.

1.INTRODUCTION

Rainfall plays a major role in indian economy.It causes lots of problem in many sector.

Our project goal is to develop a Rainfall prediction model using machine learning using a simple algorithm of neural networks.We use (ARR)Artificial Neural Network,where the research says that it gives a 90% 0f accurate result before feature selection and 91% after

feature selection.Rainfall prediction is a common application of machine learning that uses historical data and other and other factors to predict the amount of rainfall.

Rainfall influenced by many factors,including temperature,humidity,wind speed,and atmospheric pressure.These factors are interconnected and create a complex system that's difficult to forecast. It is the common application of machine learning,and linear regression is a simple and effective technique that can be used for this purpose.In this task ,the goal to predict the amount of rainfall based on historical data.After training the three models we will be predicting the disease for the input symptoms by combining the prediction of all three models.This makes our overall prediction robust and accurate.

We compare the model which we designed to the previous design to get the confirmation of our result.There are lots of hardware tools and components to predict rainfall but those tools didn't work so getting updated to the model world we use machine learning to predict rainfall.

2.SOFTWARE PROJECT DEVELOPMENT

In many software applications program different methods and cases are followed such as, Waterfall model,Iterative model, Spiral model, V model and Big Bang model. I used waterfall model in this application. I tried to use test case and case software approaches

3.OVERVIEW OF THE DESIGNED PROJECT

At first, we take the dataset from out resource then we have to perform data-preprocessing, visualization methods for cleaning and visualizing the dataset respectively and we applied the Machine Learning algorithms on the dataset then we generate the pickle file for best algorithm and flask is used as user interface for displaying the result.

4.PROPOSED SYSTEM

- **Exploratory Data Analysis of Rainfall Prediction:**Multiple datasets from different sources would be combined to form a generalized dataset, and then different machine learning algorithms would be applied to
  extract patterns and to obtain results with maximum accuracy.
- **Data Wrangling:**In this section of the report will load in the data, check for cleanliness, and then trim and clean given dataset for analysis. Make sure that the 7 document steps carefully and justify for cleaning decisions.
- **Data collection:** The data set collected for predicting given data is split into Training set and Test set. Generally, we split the dataset into Training set and Test set. The Data Model which was created using machine learning algorithms are applied on the Training set and based on the test result accuracy, Test set prediction is done.
- **Building the classification model:** For predicting the rainfall, ML algorithm prediction model is effective because of the following reasons: It provides better results in classification problem. It is strong in preprocessing outliers, irrelevant variables, and a mix of continuous, categorical and discrete variables. It produces out of bag estimate error which has proven to be unbiased in many tests and it is relatively easy to tune with.

5.FLOW DIAGRAM

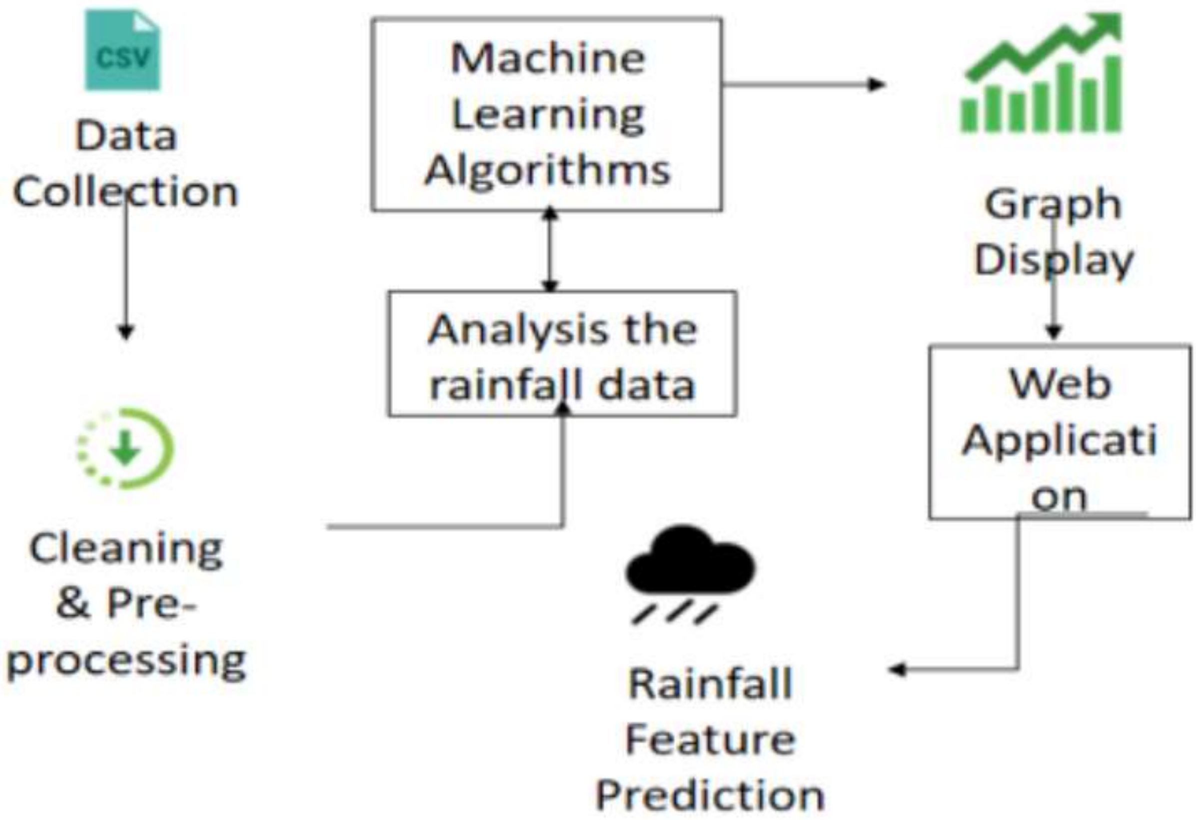


FIG: 1.FLOW CHART DIAGRAM

5.1.WORK FLOW IMPLEMENTATION:

- Initially we will need the Development Environment, for that we use Anaconda Python which has the pre-requirements like Python 3, Jupyter Notebook, Spyder.Many libraries of python are to be used like numpy, pandas ,matplotlib, scipy.
- Collect the history of the datasets required for the comparisons.
- Next the pre-processing will happen before the execution of the project.
- Missing values are a common issuse in machine learning.This occurs when a particular variable lacks data points,resulting in incomplete information and potentially harming the accuracy.
- Data feature reduction.
- Training the data.
- Testing the data.
- Predicting the model which was design.
- Result will be taken.

6. REQUIREMENT ANALYSIS

6.1. FUNCTIONAL REQUIREMENTS:

Initially we will need the Development Environment, for that we use Anaconda Python which has the pre-requirements like Python 3, Jupyter Notebook, Spyder.Many libraries of python are to be used like numpy, pandas , matplotlib , scipy.

6.2. NON-FUNCTIONAL REQUIREMENTS:

To a rainfall prediction model using linear regression and lstm algorithms.

6.3. SOFTWARE REQUIREMENTS:

pre-requirements like Python 3, Jupyter Notebook, Spyder. Many libraries of python are to be used like numpy, pandas, matplotlib , scipy.

6.4. TECHNOLOGIES USED:

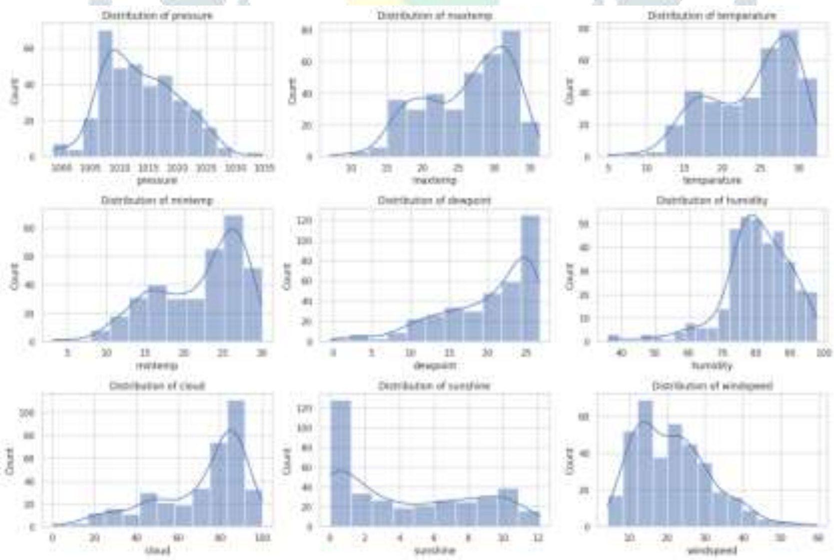
CSS: Styles the form and the displayed data,making the application visually appealing

Python: General purpose open-source programming language

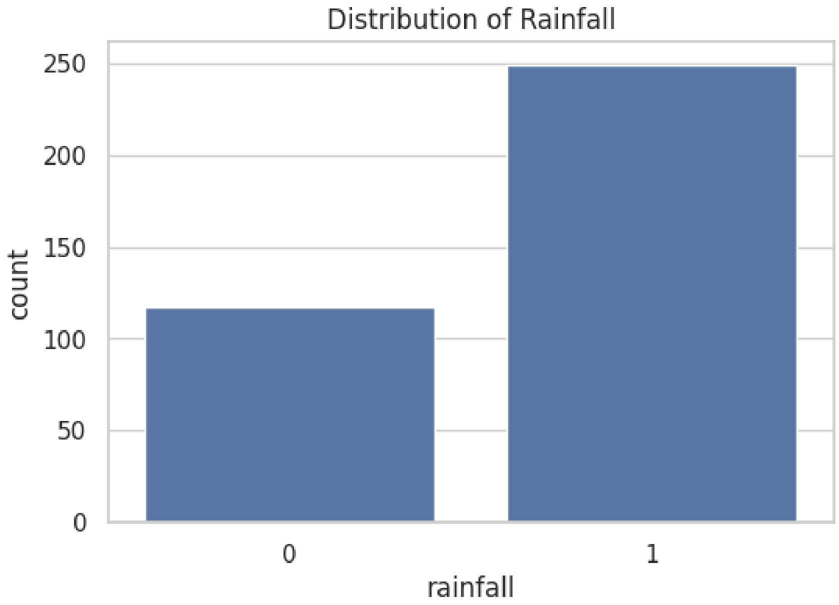
PHP: Handles server-side processing and displays the submitted data in a formatted manner

7. TESTING

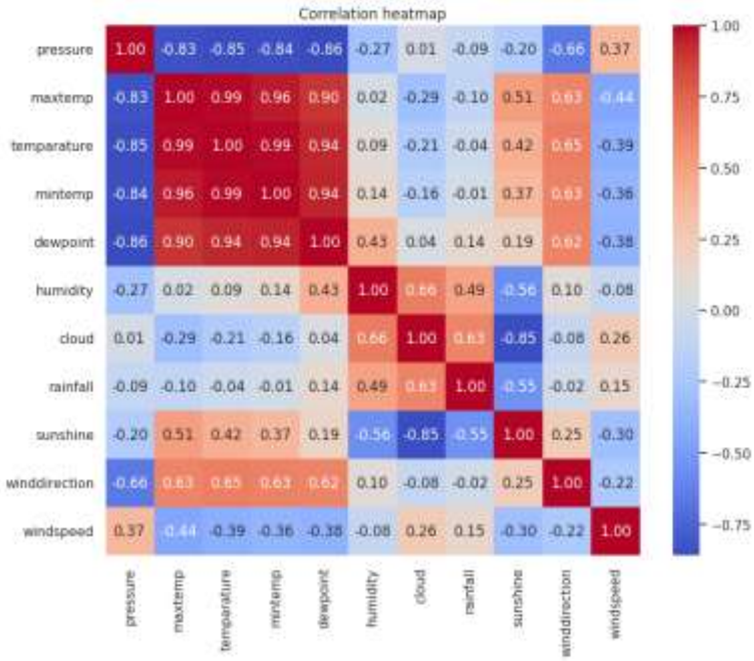
7.1:DISTRIBUTION OF COLUMN:



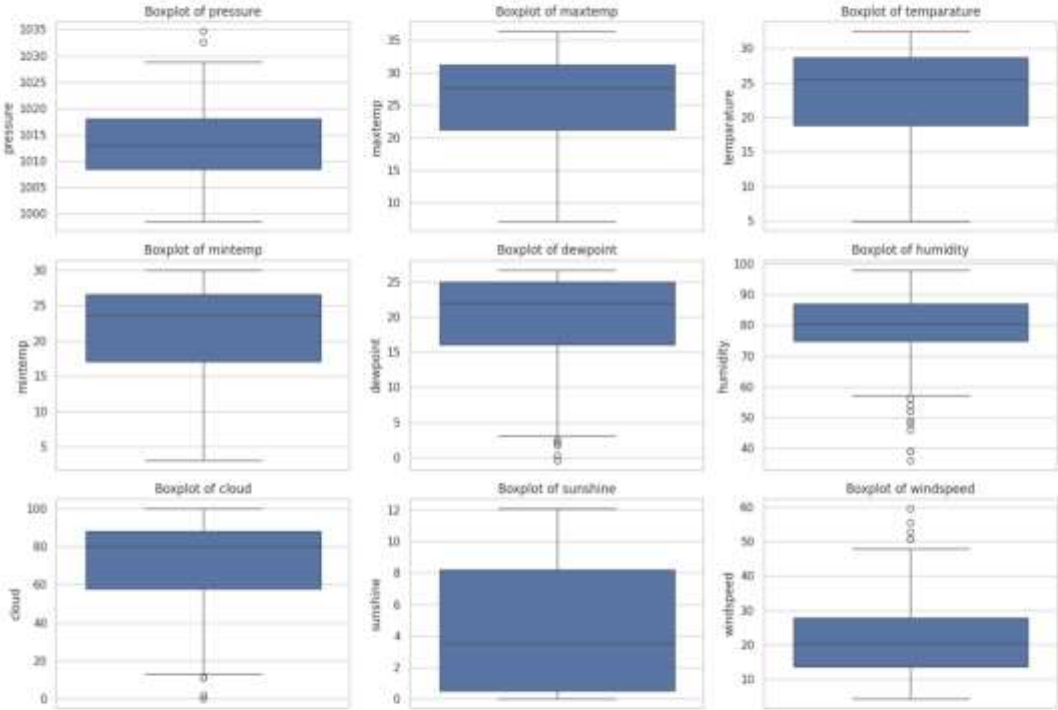
7.2:Distribution of Rainfall:



7.3:Corelation of heatmap:

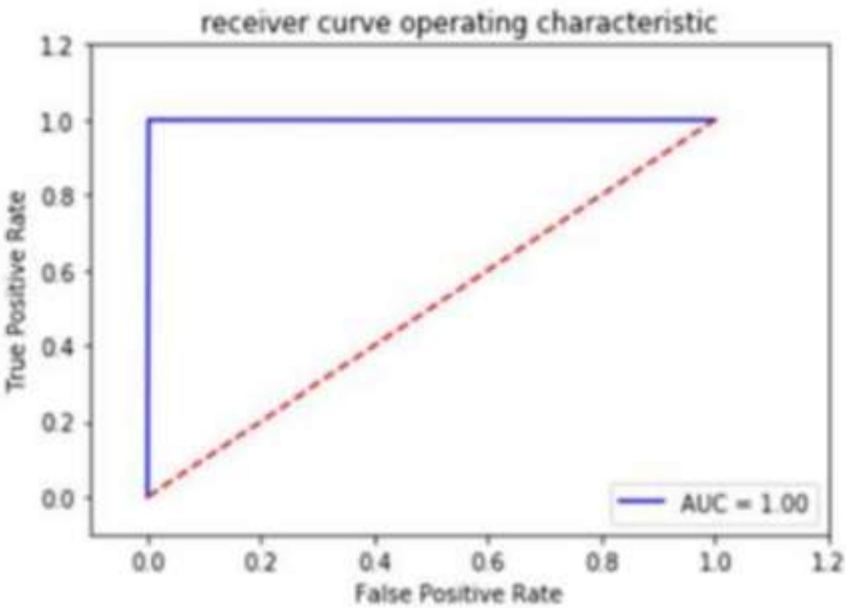


7.4:boxplot of column:



8.RESULTS AND DISCUSSION

This section discusses the implications for the proposed MLR rain-based prediction method. Total amount of data in The selected data set is 4116. Figure describes the MLR predictive effect. The accuracy of the MLR prediction is 0.99 and shown in Figure . Comparison of performance parameters are shown in Table



While discussions provide avenues for exploration and discovery, leading a discussion can be anxiety-producing: discussions are, by their nature, unpredictable, and require us as instructors to surrender a certain degree of control over the flow of information. Fortunately, careful planning can help us ensure that discussions are lively without being chaotic and exploratory without losing focus. When planning a discussion, it is helpful to consider not only cognitive, but also social/emotional, and physical factors that can either foster or inhibit the productive exchange of ideas.

9.CONCLUSION & FUTURE SCOPE

In this project we will be developing a rainfall prediction model by training and testing the dataset on linear regression and LSTM algorithms.The accuracy of algorithms will be compared and the results are obtained. It uses multi-backlash and provides improvement results depending an accuracy,MSE and corelation.
The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test set is higher accuracy score will be find out. This application can helps in predicting the Rainfall.

9.1.FUTURE SCOPE :

A weather forecast is a forecast of a particular place, using scientific and technological know-how. These include temperature, rainfall, clouds, wind speed, and humidity. Weather alerts are a special type of short-range weather forecast designed to protect human health A weather forecast is a forecast of a particular climate using the use of science technology.

These include temperature, rainfall, clouds, wind speed, and humidity. Weather alerts are a special type of shortrange weather forecast designed to protect human health.
The scope of this paper is to implement and investigate how different supervised binary classification methods impact default prediction. The model evaluation techniques used in this project are limited to precision, sensitivity, F1-score.