RAINFALL PREDICTION

By Sri Kavi Selvan & Karen Joseph

AI Model for Understanding Rainfall pattern

CLIMATE ACTIONS

Accurate rainfall predictions can help policymakers understand and mitigate the impacts of climate change, such as droughts, floods, and wildfires. This can contribute to achieving climate action by informing evidence-based policies and strategies to reduce greenhouse gas emissions and increase resilience to climate change.

DATASET USED

https://github.com/URK21CO3009KARENJOSEPH/Climate-Action/blob/main/Dataset.csv

GIHUB REPOSITORY

https://github.com/URK21CO3009KARENJOSEPH/Climate-Action/blob/main/Rainfall.ipynb

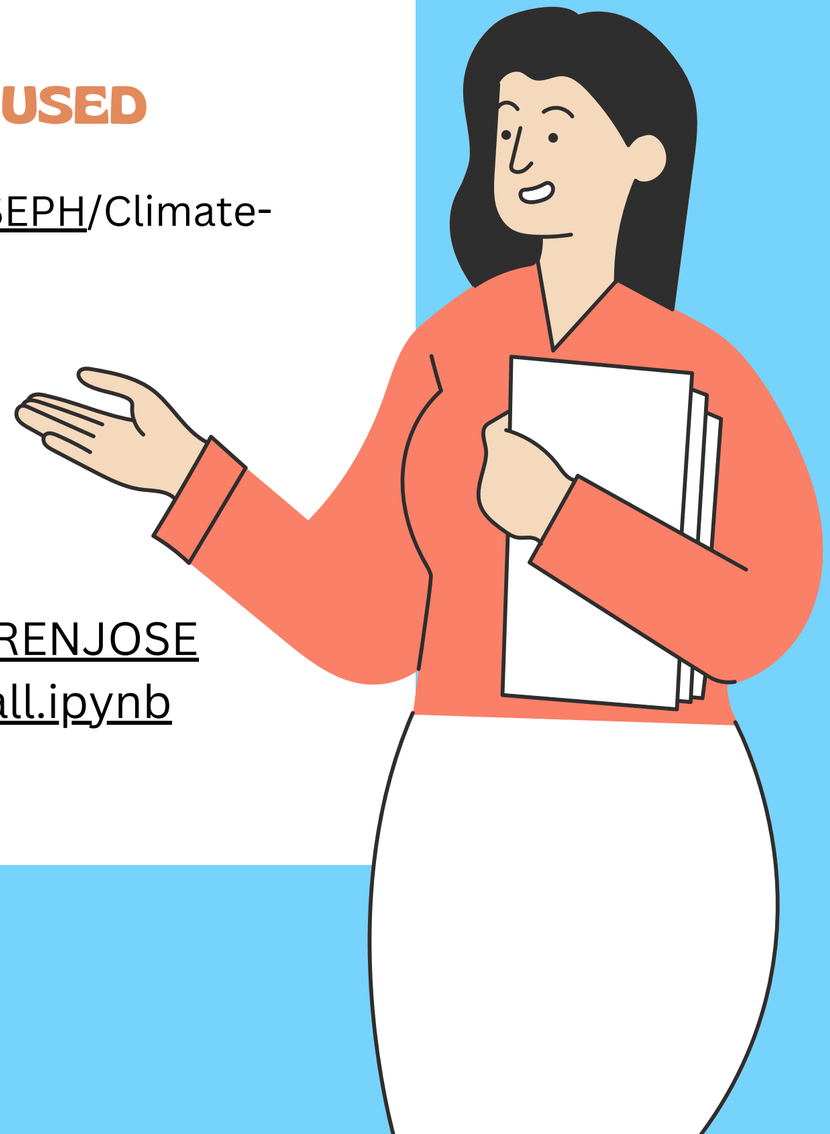


TABLE OF CONTENTS

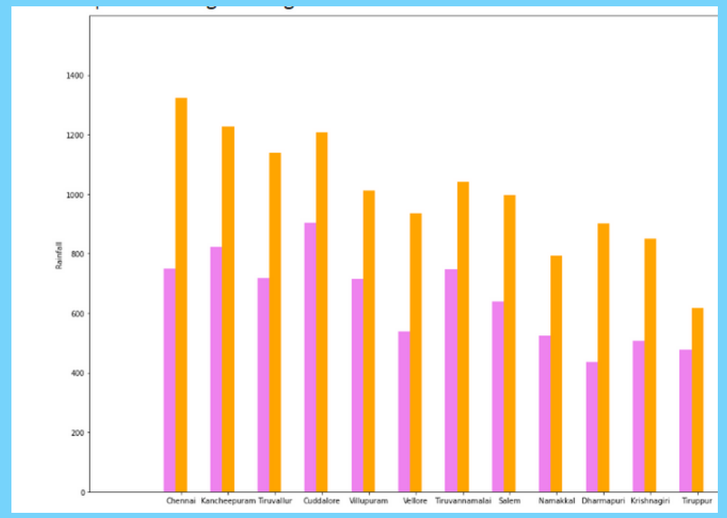
PROBLEM STATEMENT

The problem statement is to create a model that can analyze district-wise rainfall data and help understand the pattern of rainfall. The model would likely involve statistical analysis and machine learning techniques to identify trends and patterns in the data, such as seasonal or geographic variations in rainfall. The goal is to gain insights into the rainfall patterns, which could be useful for agriculture, water resource management, and disaster preparedness planning.

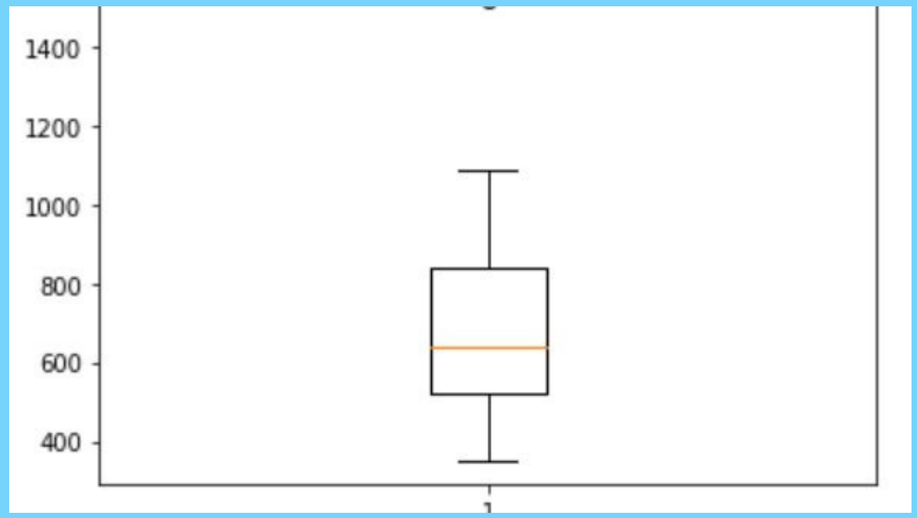
SOLUTIONS

1. Early warning systems
2. Clean Water and Sanitation.
3. Sustainable Cities and Communities
4. Agricultural planning

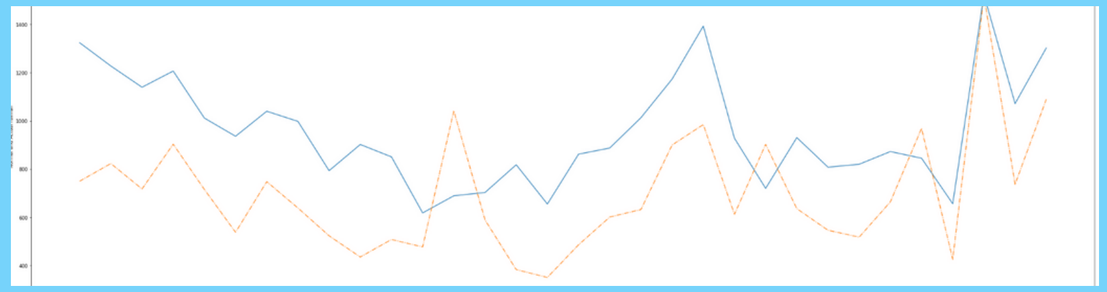
DATA VISUALISATION



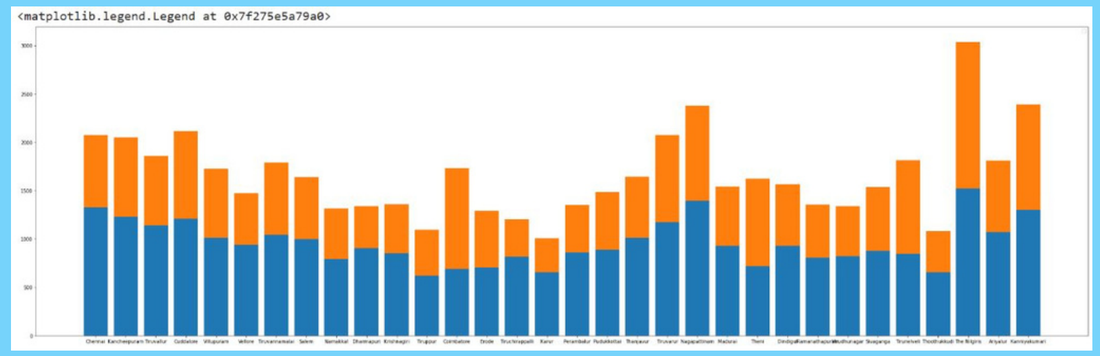
Comparative bar chart



BOX plot



Comparative Line plot



Stacked Bar chart

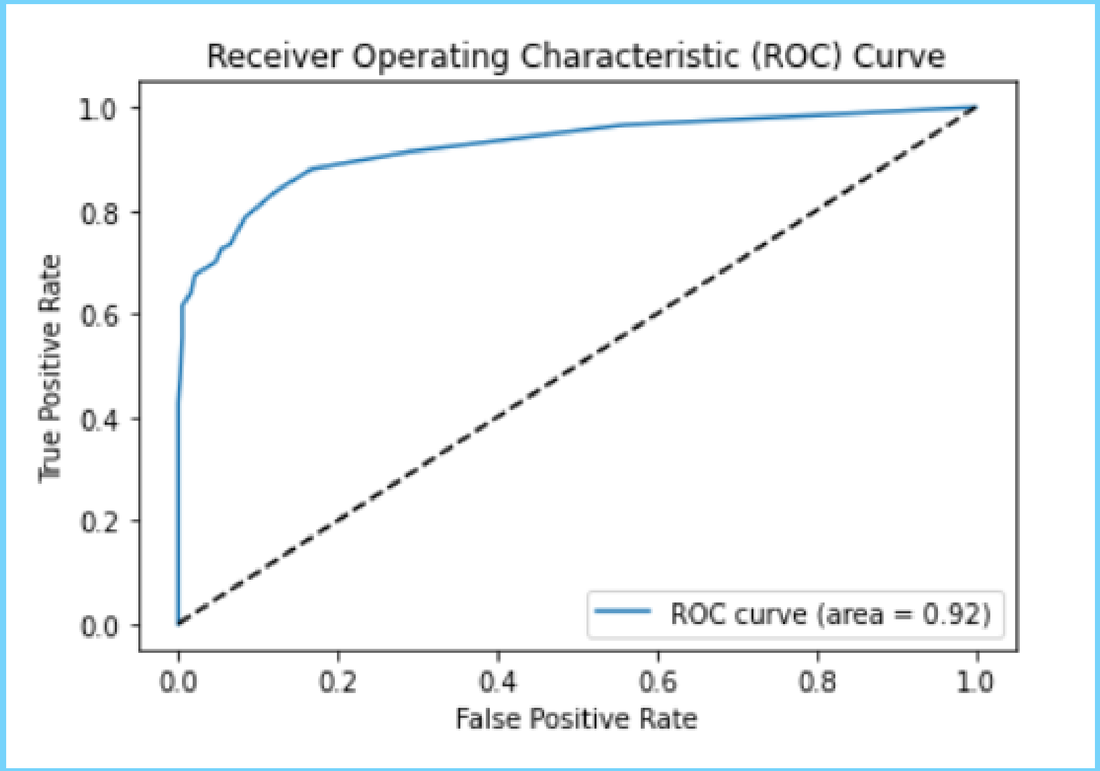
MODEL BUILDING

KNN models are trained on a labeled dataset, where each data point has a known class or value. The model predicts the label of a new data point by comparing it to the k-nearest neighbors in the training set.

Tools and Alorigthms used

- KNN Alogrithm
- Linear Regression
- Matplotlib/ Seaborn
- Skit Learn
- Numpy/ Pandas

PERFORMANCE METRICS




```
Confusion Matrix:
 [[377   6]
 [ 42  75]]
Accuracy:  0.904
Specificity:  0.9843342036553525
Recall:  0.6410256410256411
Precision:  0.9259259259259259
F1-score:  0.7575757575757577
Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.98      0.94       383
           1       0.93      0.64      0.76       117

    accuracy                           0.90       500
   macro avg       0.91      0.81      0.85       500
weighted avg       0.91      0.90      0.90       500

AUC score:  0.9241146147151369
True positive:  75
True negative:  377
```

```
Mean absolute error: 0.096
Mean squared error: 0.096
variance score: 0.46441721898640964
Root Mean Squared Error: 0.30983866769659335
```

PERFORMANCE METRICS

- The ROC curve helps to visualize the trade-off between TPR and FPR for different classification thresholds, and it can be used to evaluate the overall performance of a binary classification model. The area under the ROC curve (AUC) is a common metric used to quantify the performance of the model, with a higher AUC indicating better model performance.
- The confusion matrix will help to find the how much the model found the true positive and true negative
- The mean absolute error, mean squared error, variance score, RMSE shows that the model is accurate
- Also the accuracy is 0.904 which is 90%
- Since the performance of the model is analyzed.

INSIGHTS AND REFERENCES USED

- kaggle
- Pandas Documentations
- KNN Documentations
- Rainfall prediction using k-NN algorithm by M. R. Islam, M. A. Islam, and M. A. Hossain

THANK YOU!

Have a great day ahead.