Tinkerers Internship

Task-5: Making a prediction model

Name: Avin Shejwal

Roll No.: 52

Class: D6A

DOS: 05/07/2021

Problem Statement

Floods are among the most common disasters and natural hazard in the world, affecting human lives and causing severe economic damage. It is understood that flood risks will not decrease in the future and with the beginning of climate change, flood intensity and frequency will threaten many regions of the world. To minimize the extent of damages caused by flood, warning systems to inform the people of the disaster should be implemented in high-risk areas. This system will be able to reduce the damages of flood.

The southern most state of Kerala experienced once in a century flood.There was a huge damage to life and property. So the system should be design to be able to detect and estimate the disaster.

Machine Learning algorithms use to predict the chances of Flood in the state of Kerala using the Kerela flood dataset.

Our Approach

- A dataset with the amount of rainfall and the occurrence of flood in a particular area/state/city, in the previous years, is used. The dataset have the rainfall data from 1901 to 2018.

- Using this dataset, the average rainfall of June, July, August and September is taken and plotted it on a graph to visualize it. The average data of rainfall is taken as an input to our machine learning model and if it causes a flood or not as the output labels is determined.

- Given the input data, for consecutive 5 years, we have given this data as an input, and let the model predict, if whether there is a possibility of flooding or not, by setting some threshold in the training data.

- This Model uses 5 Machine Learning Algorithms namely KNN Classification, Logistic Regression, Support Vector Machine, Decision Tree and Random Forest to get the best possible model to predict the floods using Kerala Rainfall Data.

Platform & Code

Google colab notebook is used to write a clean and error-free python code.

Steps followed in colab notebook

- Importing Required Libraries
- Finding number of missing values
- Analyzing the Dataset
- Plotting the Data
- Using sklearn to develop the ML Model
- Using Prediction Algorithms:
  - ➢ KNN classifier to predict
  - ➢ Logistic Regression to predict
  - ➢ Support Vector Classification to predict
  - ➢ Decision Tree Classification to Predict

➤ Random Forest Classifier to predict

- Final Accuracy of our Models(Comparing the above mentioned prediction model).

Conclusion

The maximum accuracy of 0.95834 is obtained from Support Vector Classification Model which makes it a best model in our for Prediction.
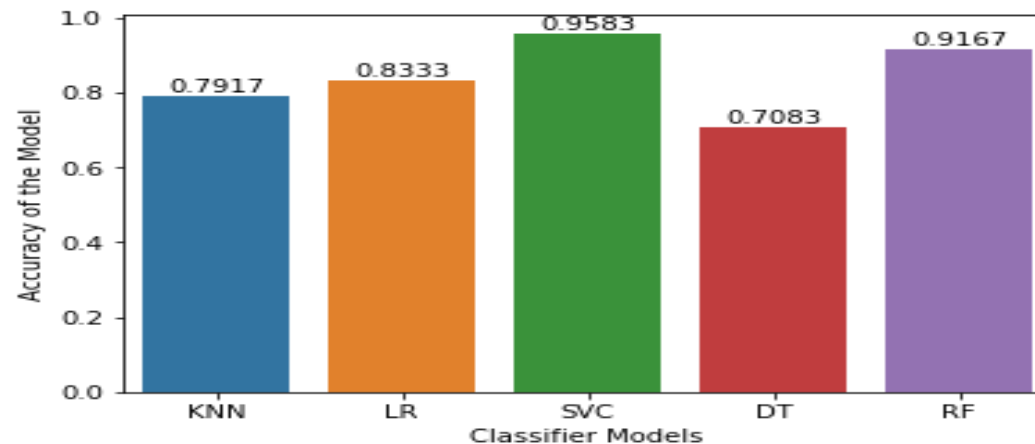
We have predicted floods just for the state of kerala depending on the rainfall data. However this method can be used for prediction for any state of india, with the given data.