**PROJECT REPORT ON**

**FLOOD PREDICTION USING IBM WATSON**

**Submitted by**

Saurish Sharma

**Under the guidance of:**

Shivani Kapoor

and

Mahidhar Chinthamreddy

**June, 2022**



**INTRODUCTION**

1.1 Overview

Floods are a major cause of loss of lives, destruction of infrastructure, and massive damage to a country’s economy. Floods, being natural disasters, cannot be prevented completely; therefore, precautionary measures must be taken by the government, concerned organizations such as the United Nations Office for Disaster Risk Reduction and Office for the coordination of Human Affairs, and the community to control its disastrous effects. To minimize hazards and to provide an emergency response at the time of natural calamity, various measures must be taken by the disaster management authorities before the flood incident. This involves the use of the latest cutting-edge technologies which predict the occurrence of disaster as early as possible such that proper response strategies can be adopted before the disaster.

Flood forecasting is the use of forecasted precipitation and streamflow data in rainfall-runoff and streamflow routing models to forecast flow rates and water levels for periods ranging from a few hours to days ahead, depending on the size of the watershed or river basin. Flood forecasting can also make use of forecasts of precipitation in an attempt to extend the lead-time available. Flood forecasting is an important component of flood warning, where the distinction between the two is that the outcome of flood forecasting is a set of forecast time-profiles of channel flows or river levels at various locations, while "flood warning" is the task of making use of these forecasts to tell decisions on warnings of floods. Real-time flood forecasting at regional areas can be done within seconds by using the technology of artificial neural networks. Effective real-time flood forecasting models could be useful for early warning and disaster prevention.

1.2 Purpose

Flood forecasting is needed for developing appropriate measures to control flood risk, mitigate flood hazard, evacuate people from flood hazard areas and manage environmental and water resources systems. This will help take preventive measures for flooding disasters to reduce the overall damage caused.

Flood forecasting allows flood control managers to predict, with a high degree of accuracy, when local flooding is likely to take place.

**LITERATURE SURVEY**

2.1 Existing problem

A classification framework is presented which classifies the remote sensing technologies being used for flood prediction into three types, which are: multispectral, radar, and light detection and ranging (LIDAR). Further categorization is performed based on the method used for data analysis. The technologies are examined based on their relevance to flood prediction, flood risk assessment, and hazard analysis. Some gaps and limitations present in each of the reviewed technologies have been identified.

2.2 Proposed solution

A flood prediction and extent mapping model are then proposed to overcome the current gaps. The compiled results demonstrate the state of flood prediction whether it’s likely to be affected by flood or not.

We have created a model using machine learning techniques and by applying data science algorithms to predict Flood using particular features such as :

* Temperature
* Humidity
* Cloud Cover
* Annual
* Jan-Feb
* Mar-May
* June-September
* Oct-Dec
* AvgJune
* Sub

**THEORETICAL ANALYSIS**

3.1 Block diagram

The diagrammatic overview of the project.

3.2 Hardware / Software designing

The project is a software based project.

The softwares required for this project are:

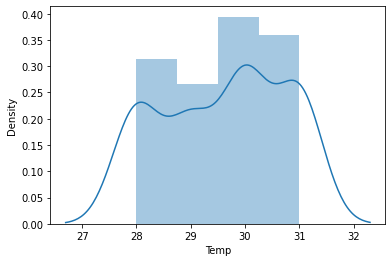
1. Jupyter Notebook
2. IBM Watson
3. VS Code

**EXPERIMENTAL INVESTIGATIONS**

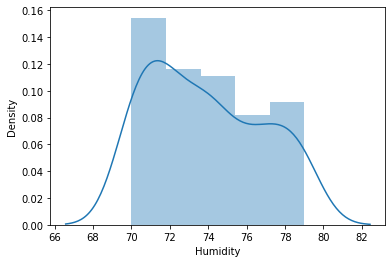
The experimental investigation includes univariate, bivariate and multivariate analysis of various independent variables given in the dataset.

**Univariate Analysis:**

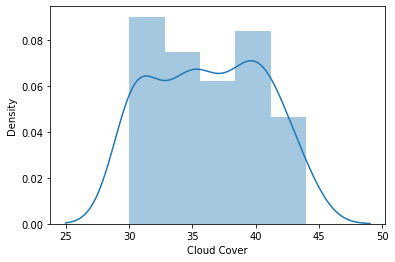
1. Temperature distribution plot



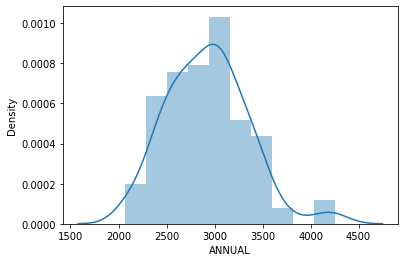
1. Humidity distribution plot



1. Cloud Cover distribution plot



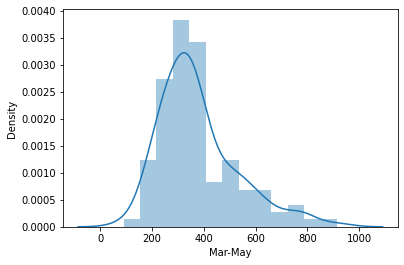
1. Annual rainfall distribution plot



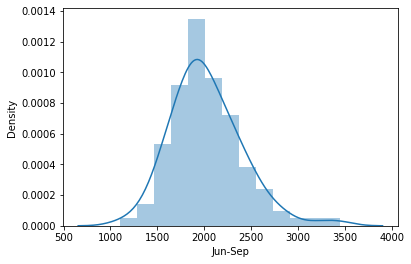
1. Jan-Feb rainfall distribution plot



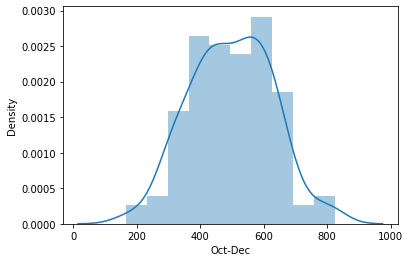
1. Mar-May rainfall distribution plot



1. Jun-Sep rainfall distribution plot

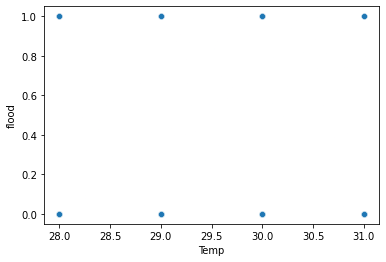


1. Oct-Dec rainfall distribution plot

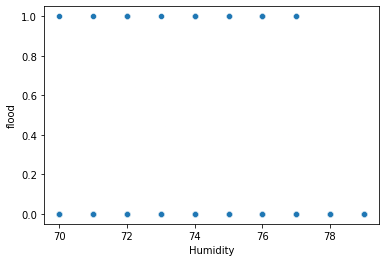


**Bivariate Analysis:**

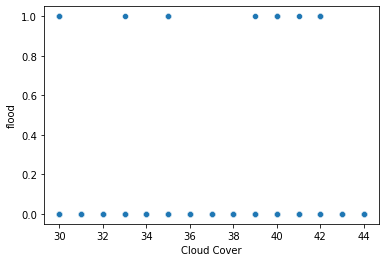
1. Temp and Flood



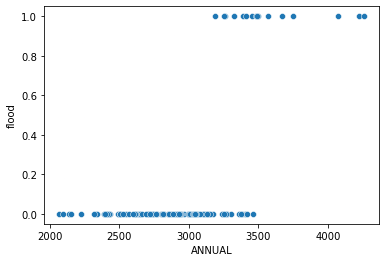
1. Humidity and Flood



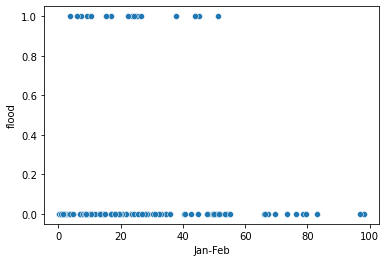
1. Cloud Cover and Flood



1. Annual and Flood



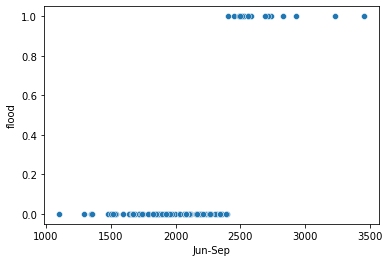
1. Jan-Feb and Flood



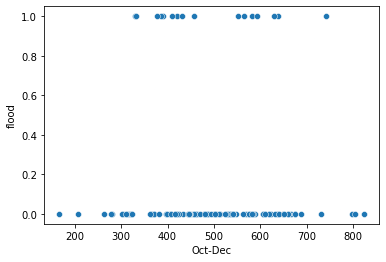
1. March-May and Flood



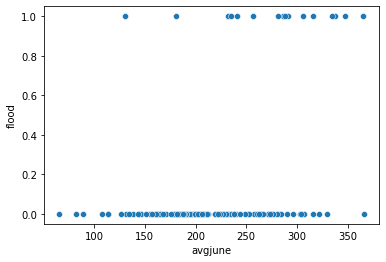
1. Jun-Sept and Flood



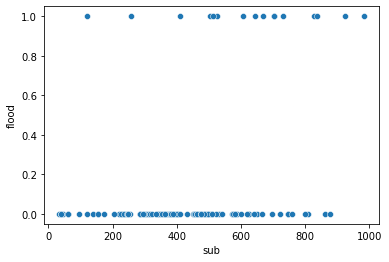
1. Oct- Dec and Flood



1. Avgjune and Flood

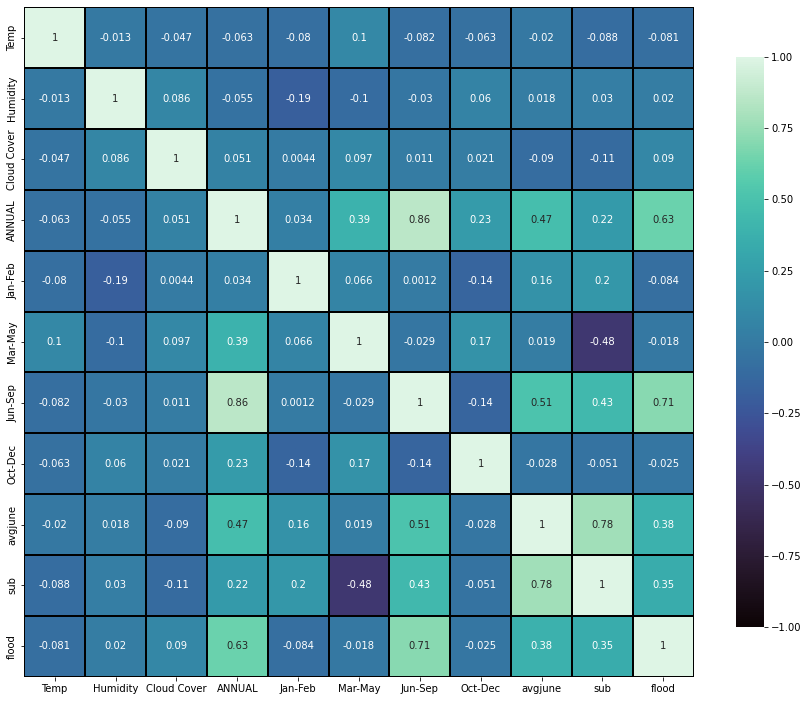


1. Sub and Flood



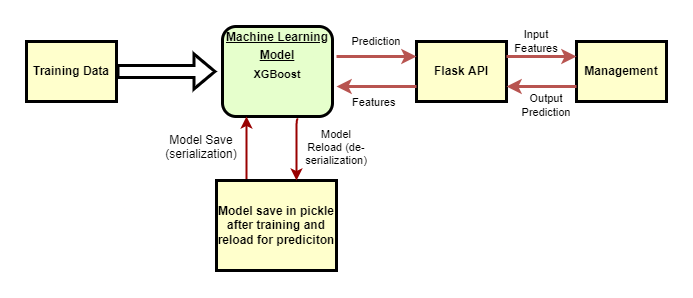
**Multivariate:**

1. Heatmap showing correlation between columns



**FLOWCHART**

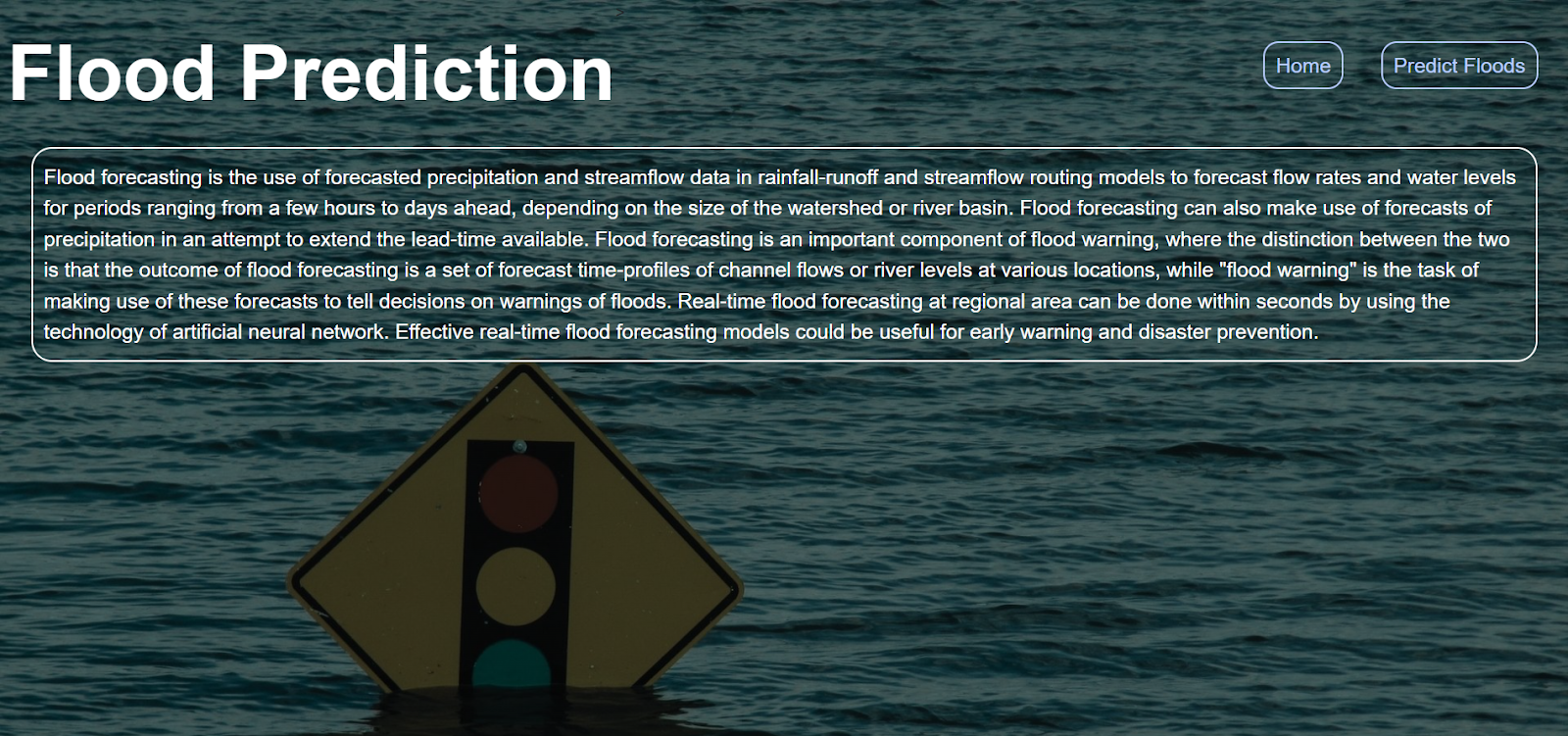
Control flow of the project



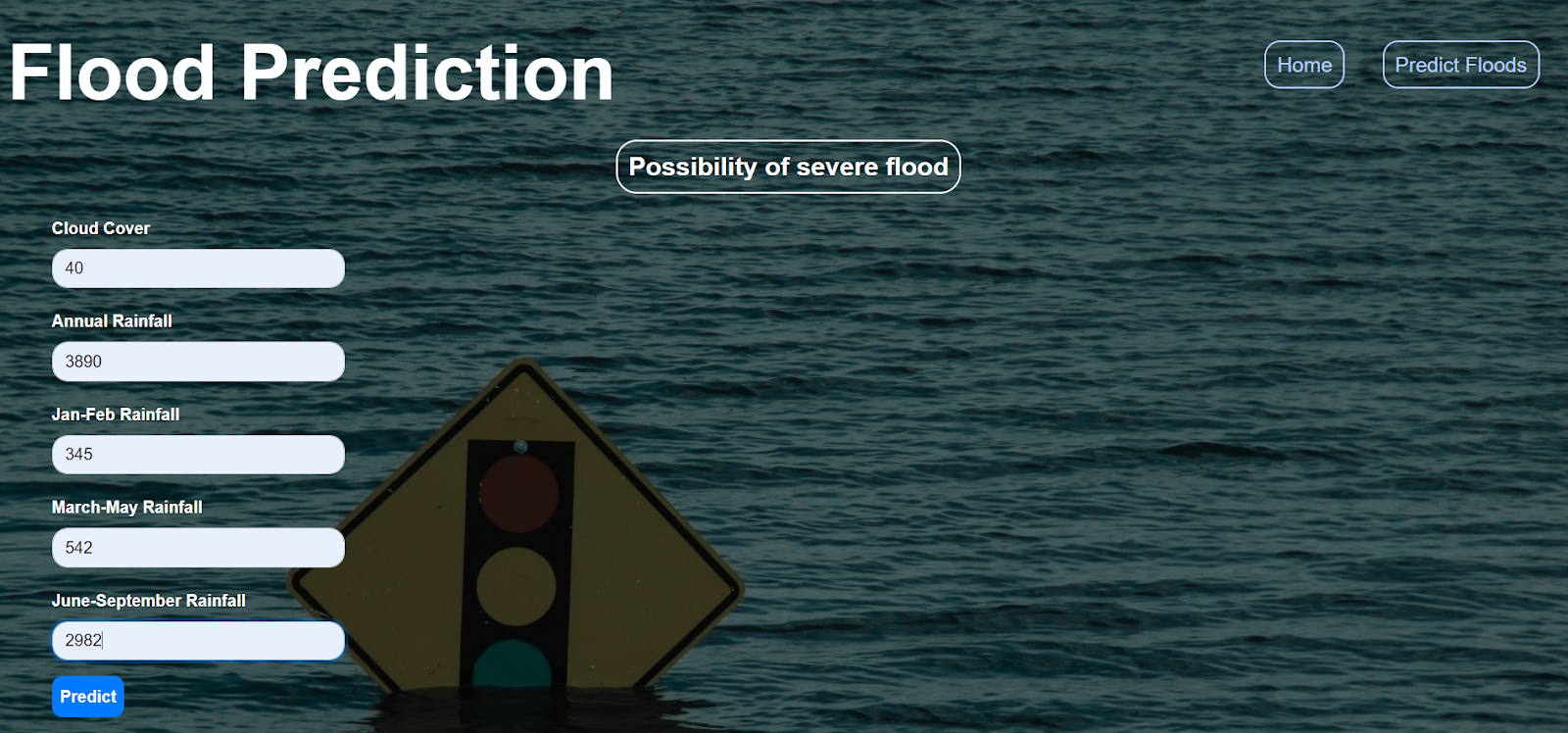
**RESULT**

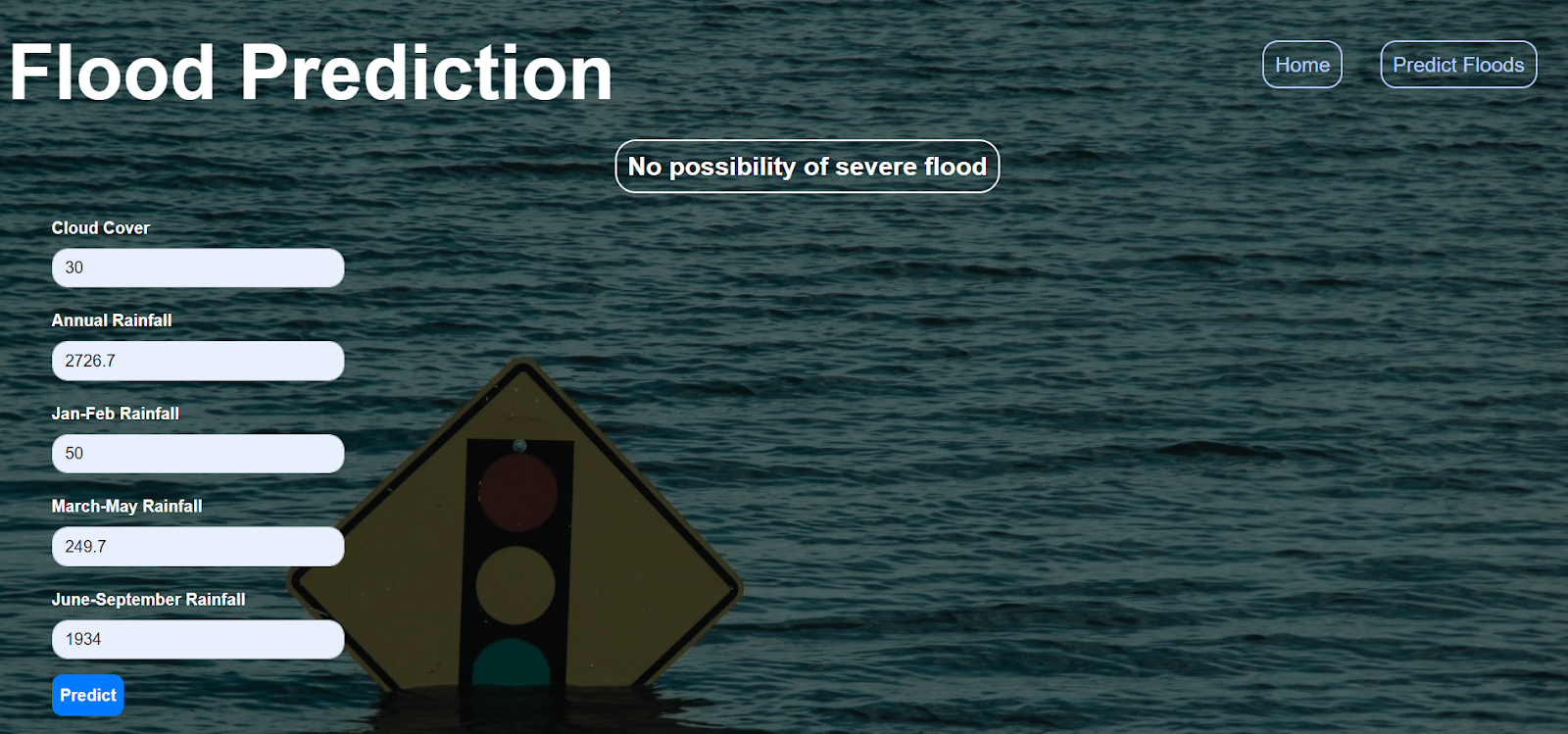
The project was made and implemented with a web application which was made using flask. The web application has two pages, the first is the homepage and the other is the page to predict the flood possibility.

1. Homepage



1. Flood Prediction





**ADVANTAGES & DISADVANTAGES**

Advantage:

1. As this model is based on machine learning, it’s already capable of predicting the flood using train and test values provided by the input dataset so we need not worry about the past issues rather just input the current status of the area to get the desired result.

Disadvantage:

1. The model has accuracy around 86% which can be improved upon in future works.

**APPLICATIONS**

Flood forecasting is needed for developing appropriate measures to control flood risk, mitigate flood hazard, evacuate people from flood hazard areas, determine insurance premiums, and manage environmental and water resources systems.

It can be applied in all the weather forecasting observatories, flood reporting observatories, meteorological stations and geological observatories for predicting floods beforehand.

It can also be set up in towns and cities that are vulnerable to floods so that they can prepare for the calamity and take all the precautions.

**CONCLUSION**

The project flood prediction using machine learning was completed successfully. The machine learning model was selected to be Xgboost among the models of Decision Tree, Random Forest, K Nearest Neighbors and Xgboost. Then the flask web application was built and both were deployed on IBM Watson cloud platform successfully.

**FUTURE SCOPE**

Alongside A.I , Machine Learning and Data Science we can also add Deep Learning, CNN(Convolutional Neural Networks - A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.) to make the flood prediction model even more accurate and efficient.

**BIBLIOGRAPHY**

* <https://www.kaggle.com/datasets/arbethi/rainfall-dataset?select=flood+dataset.xlsx>
* <https://towardsdatascience.com/top-10-python-libraries-for-data-science-cd82294ec266>
* <https://youtu.be/QeKshry8pWQ>
* <https://youtu.be/D6gtZrsYi6c>
* <https://youtu.be/aWAnNHXIKww>
* <https://youtu.be/lj4I_CvBnt0>
* <https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/>
* <https://www.w3schools.com/html/>
* <https://www.w3schools.com/css/>
* <https://youtu.be/TysuP3KgSzc>

**APPENDIX**

**APPENDIX:**   
Source Code





