FLOOD PREDECTION

Using Ibm Watson Machine Learning

Developed By: P.Avinash, A.Poojitha, K.Sruthi, K.Sai Sudha

Smart Bridge – Mini Project Report

1. INTRODUCTION

1.1 Overview:

Among the natural disasters, floods are the most destructive, causing massive damage to human life, infrastructure, agriculture, and the socioeconomic system. Governments, therefore, are under pressure to develop reliable and accurate maps of flood risk areas and further plan for sustainable flood risk management focusing on prevention, protection, and preparedness [1]. Flood prediction models are of significant importance for hazard assessment and extreme event management.

Robust and accurate prediction contribute highly to water recourse management strategies, policy suggestions and analysis, and further evacuation modeling [2]. Thus, the importance of advanced systems for short-term and long-term prediction for flood and other hydrological events is strongly emphasized to alleviate damage [3]. However, the prediction of flood lead time and occurrence location is fundamentally complex due to the dynamic nature of climate condition. Therefore, today's major flood prediction models are mainly data-specific and involve various simplified assumptions [4]. Thus, to mimic the complex mathematical expressions of physical processes and basin behavior, such models benefit from specific techniques e.g., event-driven, empirical black box, lumped and distributed, stochastic, deterministic, continuous, and hybrids.

1.2 Purpose

Floods are among the most destructive natural disasters, which are highly complex to model. The research on the advancement of flood prediction models contributed to risk reduction, policy suggestion, minimization of the loss of human life, and reduction of the property damage associated with floods. To mimic the complex mathematical expressions of physical processes of floods, during the past two decades, machine learning (ML) methods contributed highly in the advancement of prediction systems providing better performance and cost-effective solutions. Due to the vast benefits and potential of ML, its popularity dramatically increased among hydrologists. Researchers through introducing novel ML methods and hybridizing of the existing ones aim at discovering more accurate and efficient prediction models. The main contribution of this paper is to demonstrate the state of the art of ML models in flood prediction and to give insight into the most suitable models. In this paper, the literature where ML models were benchmarked through a qualitative analysis of robustness, accuracy, effectiveness, and speed are particularly investigated to provide an extensive overview on the various ML algorithms used in the field. The performance comparison of ML models presents an in-depth understanding of the different techniques within the framework of a comprehensive evaluation and discussion. As a result, this paper introduces the most promising prediction methods for both long-term and short-term floods. Furthermore, the major trends in improving the quality of the flood prediction models are investigated. Among them, hybridization, data decomposition, algorithm ensemble, and model optimization are reported as the most effective strategies for the improvement of ML methods. This survey can be used as a guideline for hydrologists as well as climate scientists in choosing the proper ML method according to the prediction task.

2. LITERATURE SURVEY

2.1 Existing problem

Flood forecasting is highly complicated and expensive. The weather and rainfall is a factor of predicting the flood. The advanced technology uses simulations supported by physics and differential equations. The satellite images are used to get the rainfall data. In recent times rapid urbanization, global climate change and extreme rainfall have resulted in flash floods. In orthodox methods of flood forecasting, using satellite images and radar also involving mathematical equations, current weather conditions are detected. At present machine learning technologies are implemented to detect such kinds of natural disasters. The floods are predicted by considering the parameters causing the flash flood. There are some drawbacks in machine learning that lead to wrong predictions of floods. The results cannot be accurate in predicting flash floods..

2.2 Proposed solution

The aim of this project is to get all the rainfall data of India and from a dataset containing yearly rainfall data. By providing real time input to different models of machine learning, those are Logistic Regression, Support Vector Machine, K-Nearest Neighbors and Decision Tree Classifier. The input provided to models are pre-processed and patterns are extracted by getting maximum accuracy. The data provided is split into a Training set and Test set. It is split in the ratio of 7:3. The all four models are used to predict and by comparing all the results of model and considering the confusion matrix of all the models the accuracy is determined. The best model is chosen by comparing the accuracy of each modelNaïve Bayes: It is a division of "probabilistic classifiers" which is based on relating Bayes' theorem with strong independence expectations between the structures. They belong to basic Bayesian network models, however it is combined with kernel density approximation, High accuracy level can be achieved.

3.THEORITICAL ANALYSIS:

3.1 Block diagram

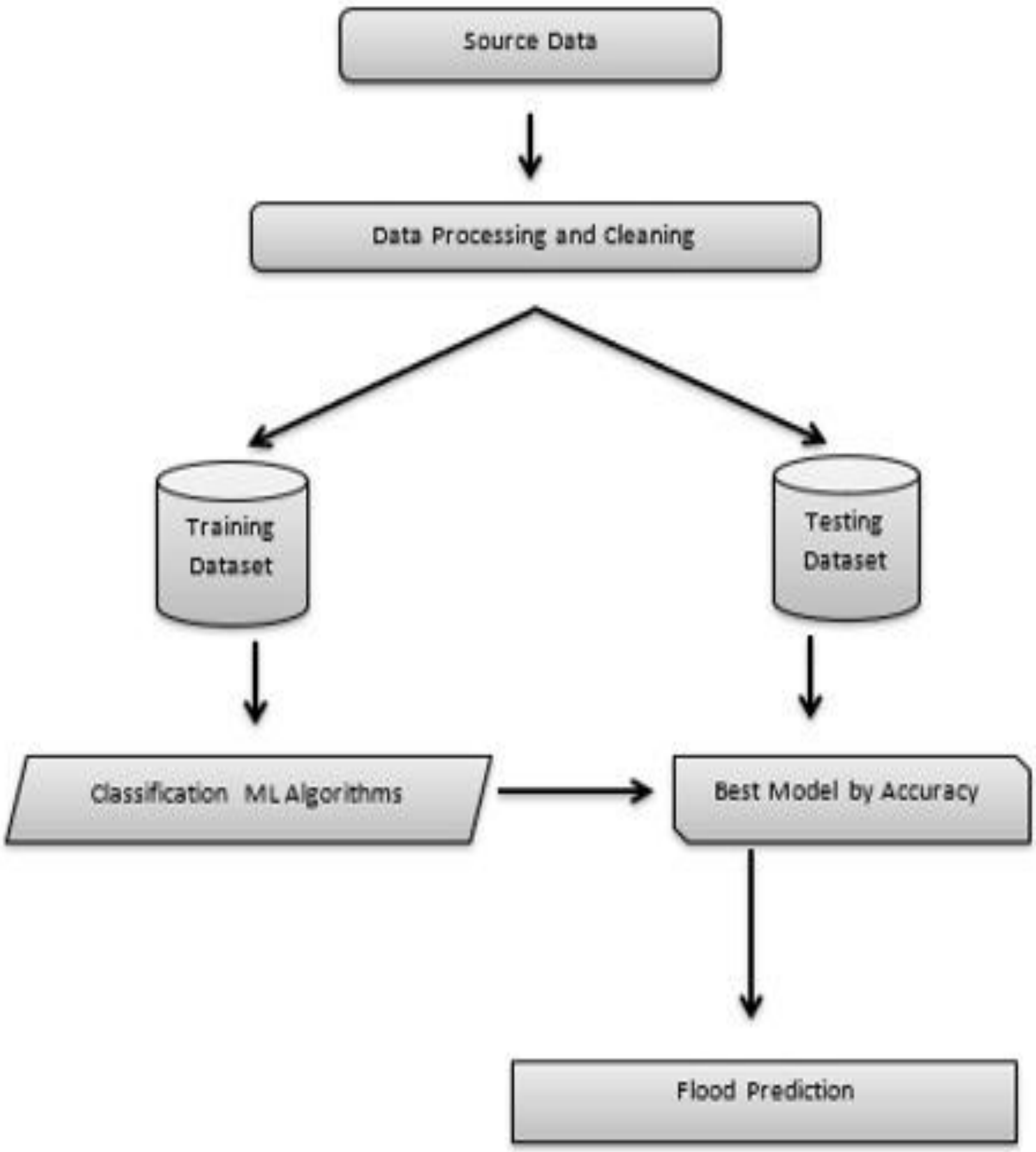


Fig (i) Block Diagram

3.2HARDWARE AND SOFTWARE REQUIREMENTS IN THE PROJECT:

For running a machine learning model on the system you need a system with minimum of 16 GB RAM in it and you require a good processor for high performance of the model.

In the list of **software requirements** you must have:

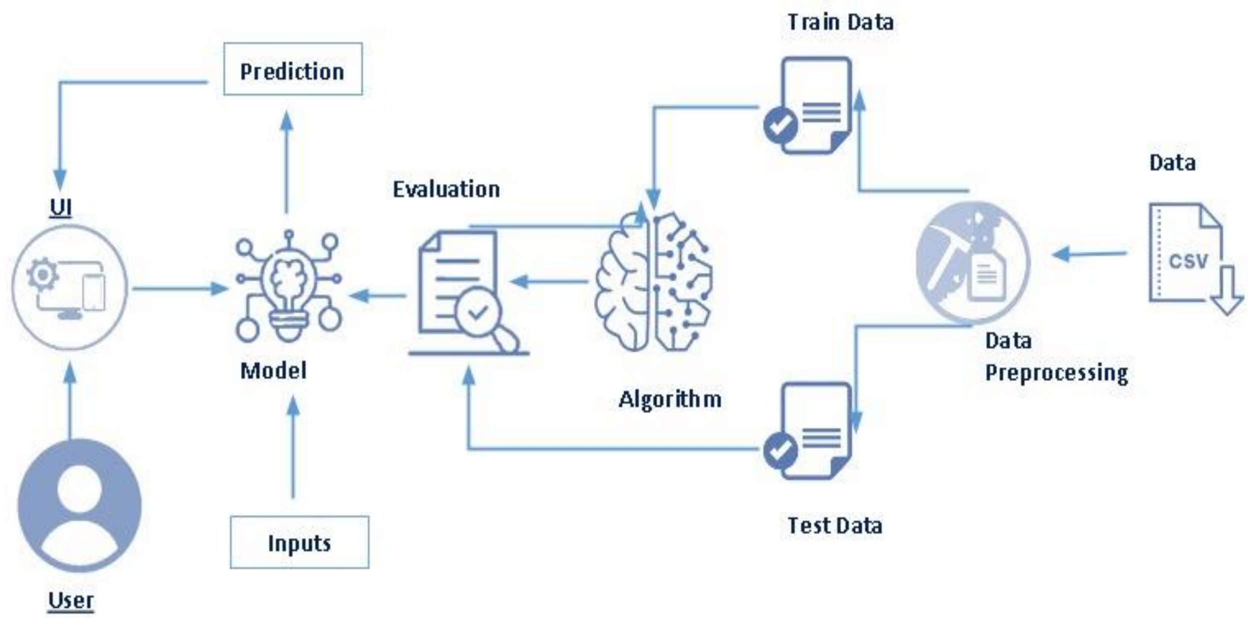
- Jupyter Notebook for programming, which can be installed by Anaconda IDE.
- Python packages
- A better software for running the html and css files for application building phase e.g. spyder.

4.EXPERIMENTAL INVESTIGATIONS:

4.1 Data Preprocessing:-

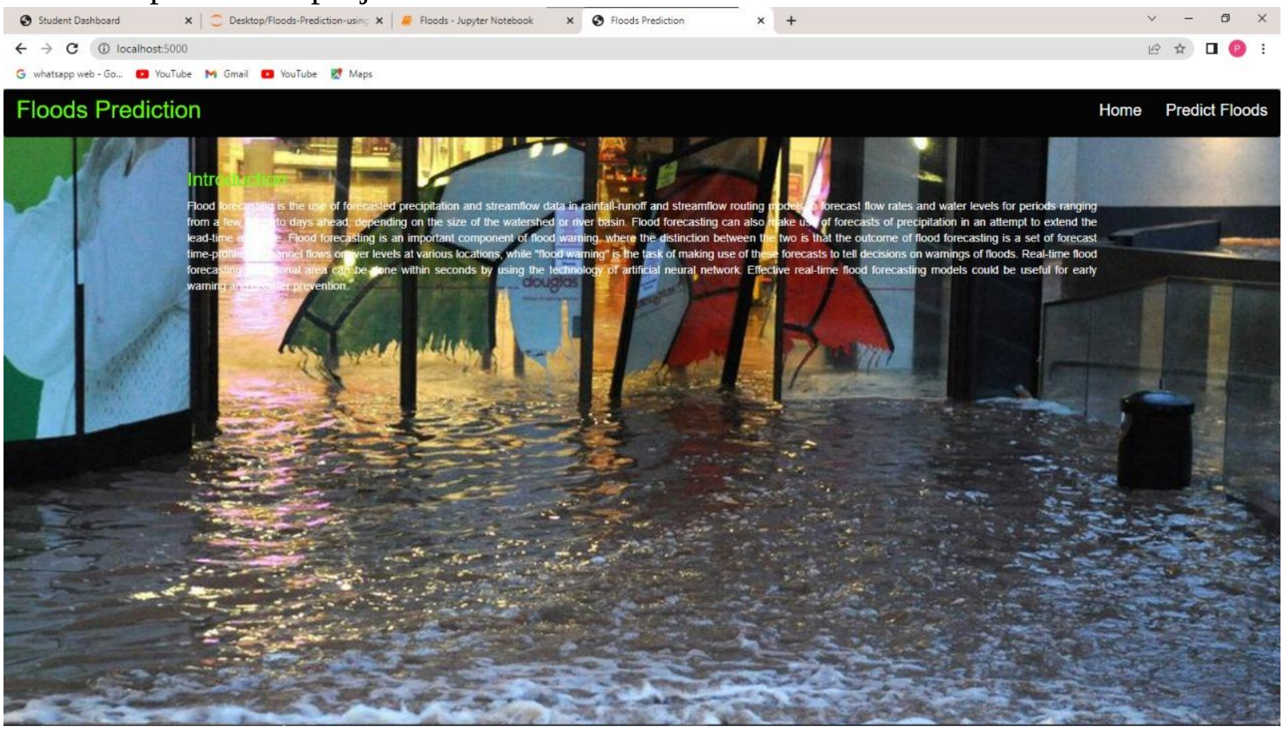
- Define a problem
- Preparing data
- Evaluating algorithms
- Predicting results
- Predicting results

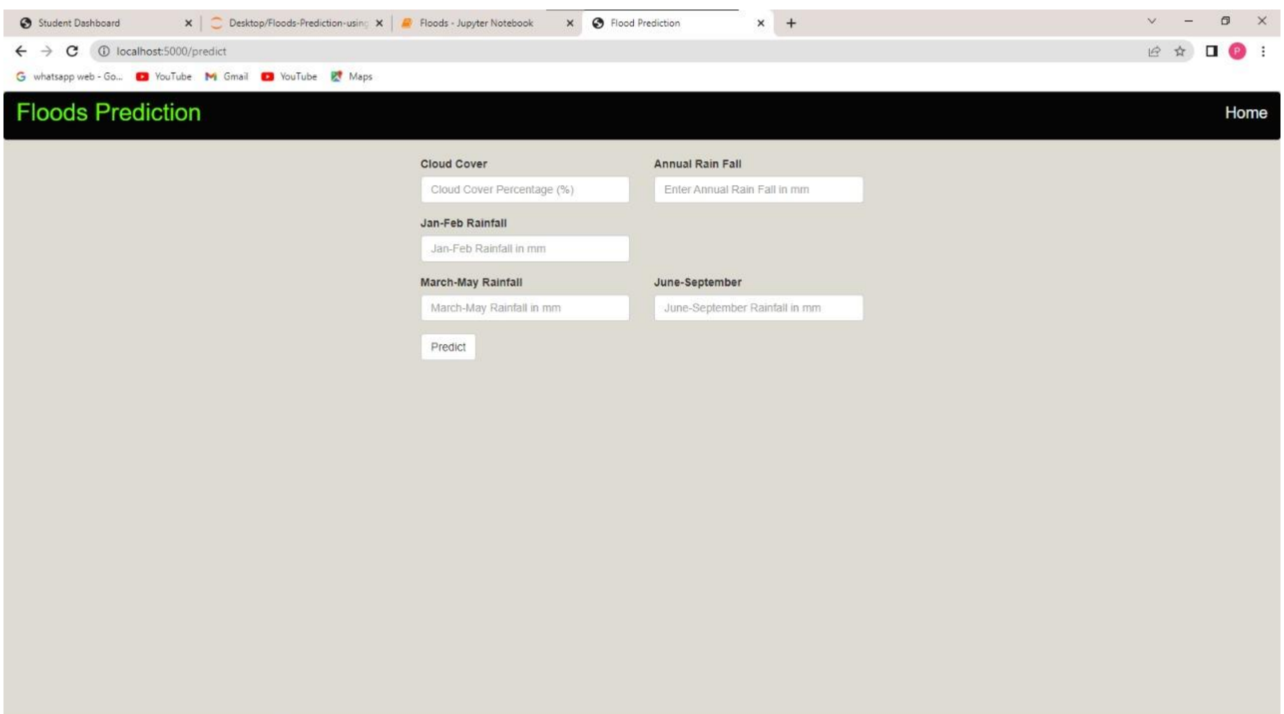
5.FLOW CHART:-

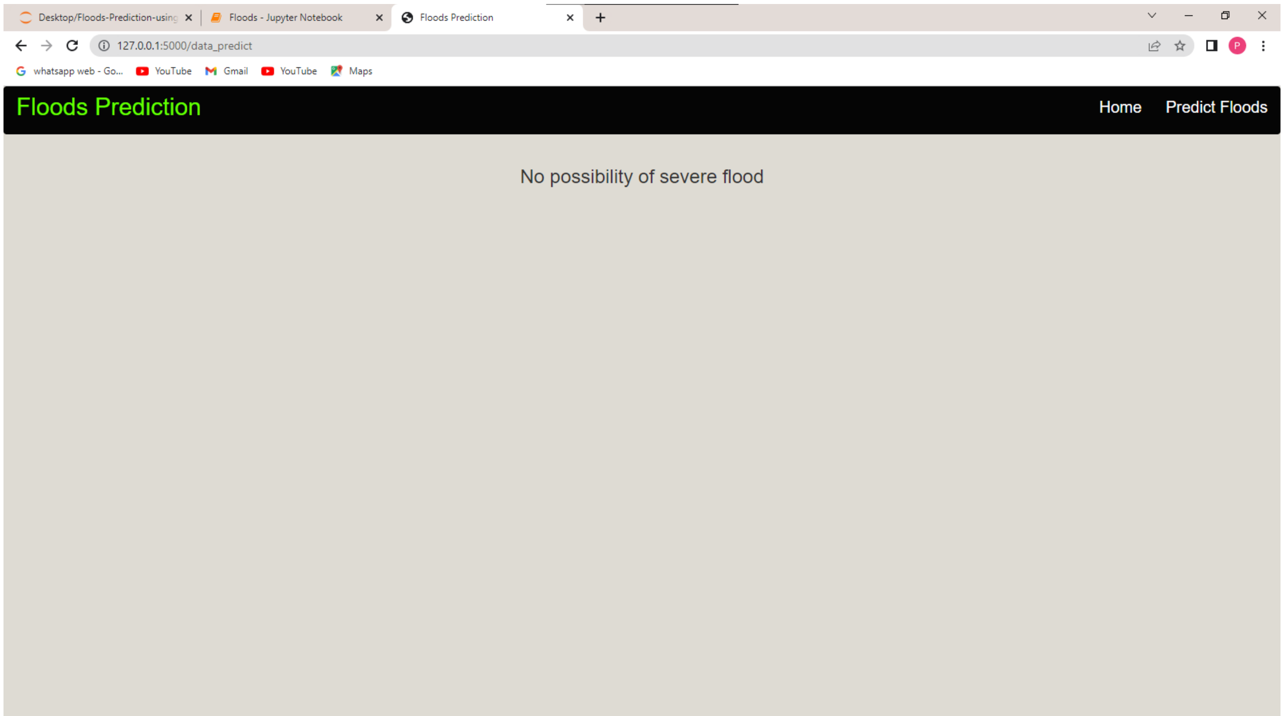


6.RESULTS:-

Final output of the project:







7. ADVANTAGES & DISADVANTAGES

ADVANTAGES

- The impact of flooding is reduced
- Warnings give people time to move possessions upstairs, put sandbags in position and to evacuate.

DISADVANTAGES

- Warnings don't stop a flood from happening
- Living in a place that gets lots of warnings could make it difficult to get insurance
- People may not hear or have access to warnings

8.APPLICATIONS:-

Flood is an overflowing of a great body of water over land not usually submerged (Taiwo et al, 2019). It occurs when there is an excess of water which in turn covers a land that is usually dry. It is also regarded as the most frequent type of natural disaster which occur when an overflow of water submerges land that is usually dry (World Health Organisation – WHO). Floods are devastating natural disasters worldwide, it is a deleterious phenomenon that induces detrimental impacts on humans, properties and environment (Mind’je et al, 2019). Of all weather-related natural disasters, floods are the most common and widespread natural severe weather event which is estimated to affect 250 million people around the world every year, also costing billions of dollars in damages (Matias, 2018). Recent floods and consequences all over the world are becoming too frequent and threat to sustainable development in human settlements (Nwigwe and Emberga, 2014). According to The National Severe Storms Laboratory (NSSL), floods kill more people each year than tornadoes, hurricanes or lightning in the United States. Between 1998-2017, floods affected more than 2 billion people worldwide (WHO, 2021).

9.CONCLUSION-

The machine learning system ignited by data cleaning and processing, replacing or removing the null values, model building and evaluation. At the end the flood prediction model has given different accuracy results from four different models. From the above results and analysis, the best algorithm for flood prediction is Logistic Regression with (99%).

10.FUTURE SCOPE:-

In the future scope, We intend to investigate prediction approach with the revised data set and employ the most accurate and relevant machine learning algorithms for detection.

Therefore a system pertaining towards deep neural network is developed where flood occurrence with highest accuracy predicted by using temperature and rainfall intensity. Network is developed where flood occurrence with highest accuracy predicted by using temperature and rainfall intensity. Based on this flood prediction, an alert of DANGER is created on the respective areas through online flood portal system which can save many lives. According to the datasets collected on each days noting of values of parameters, which is fed into KNN and DNN, therefore result indicates that DNN can be efficiently used for forecasting flood with its percentage calculation. The most recent dataset can be collected from recent Kerala flood occurrence in order to achieve best result and verification of this system.

In future, the best biologically verified topological factors can be collected in forecasting flood not only in certain restricted areas but also in overall regions. Finally, by including standard ML algorithms with chat bots alerting people quickly results for further research and development of new methods. Last but not the least though people get alerted, the evacuation should also be made easy through drones implementations.

11.BIBILOGRAPHY:-

1. Danso-Amoako, E.; Scholz, M.; Kalimeris, N.; Yang, Q.; Shao, J. Predicting dam failure risk for sustainable flood retention basins: A generic case study for the wider greater manchester area. *Comput. Environ. Urban Syst.* **2012**, *36*, 423–433. [CrossRef]
2. Xie, K.; Ozbay, K.; Zhu, Y.; Yang, H. Evacuation zone modeling under climate change: A data-driven method.
   J. Infrastruct. Syst. **2017**, *23*, 04017013. [CrossRef]
3. Pitt, M. *Learning Lessons from the 2007 Floods*; Cabinet Office: London, UK, 2008.
4. Lohani, A.K.; Goel, N.; Bhatia, K. Improving real time flood forecasting using fuzzy inference system.
   J. Hydrol. **2014**, *509*, 25–41. [CrossRef]

APPENDIX :-

app.py

```
from flask import Flask, render_template, request
# used to run/serve our application
# render_template is used for rendering the html pages
#import load from joblib to load the saved model file
from joblib import load
import pickle
app=Flask(__name__) # our flask app
```

```python
#load model file
model =pickle.load(open("model.pkl","rb"))
sc=pickle.load(open("scale.pkl","rb"))
@app.route('/') # rendering the html template
def home():
    return render_template('home.html')
@app.route('/predict') # rendering the html template
def index() :
    return render_template("index.html")


@app.route('/data_predict', methods=['POST']) # route for our prediction
def predict():
    temp = request.form['temp']
    Hum = request.form['Hum']
    db = request.form['db']
    ap = request.form['ap']
    aa1 = request.form['aa1']

    data = [[float(temp),float(Hum),float(db),float(ap),float(aa1)]]
    prediction = model.predict(sc.transform(data))
    output=prediction[0]
    if(output==0):
        return render_template('noChance.html', prediction='No possibility of severe flood')
    else:
        return render_template('chance.html', prediction='possibility of severe flood')

if __name__ == '__main__':
    app.run(debug=False)
```