VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Belagavi – 590018

A Project Report on

FLOOD PREDICTION USING AI

Submitted by

AJAY KARTHIK	4SN17CS700
CHAITANYA THEKKUNJA	4SN17CS711
KISHAN KASHYAP	4SN17CS714
SHUBHANG	4SN17CS723

In partial fulfillment of the requirements for the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE & ENGINEERING

Under the Guidance of

Mrs. GANGOTHRI S

Associate Professor

Department of Computer Science & Engineering

SRINIVAS INSTITUTE OF TECHNOLOGY

(NAAC ACCREDITED)

MANGALURU - 574143, KARNATAKA.

2020 - 2021

SRINIVAS INSTITUTE OF TECHNOLOGY

(NAAC ACCREDITED)

MANGALURU - 574143, KARNATAKA.

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

*Certified that the project work entitled "**FLOOD PREDICTION USING AI**" is a bona fide work carried out by*

AJAY KARTHIK	**4SN17CS700**
CHAITANYA THEKKUNJA	**4SN17CS711**
KISHAN KASHYAP	**4SN17CS714**
SHUBHANG	**4SN17CS723**

*in partial fulfillment for the award of **BACHELOR OF ENGINEERING** in **COMPUTER SCIENCE & ENGINEERING** of the **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the year 2020 – 2021. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the Bachelor of Engineering Degree.*

Mrs. Gangothri S **Project Guide**	**Dr. Shivakumar G. S.** **Head of the Department**	**Dr. Shrinivasa Mayya D.** **Principal**

Name of the Examiners **Signature with Date**

1.

2.

ABSTRACT

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 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.

CHAPTER 1

INTRODUCTION

Flood prediction is the use of forecasted precipitation and stream-flow data in rainfall-runoff and stream-flow 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 prediction can also make use of forecasts of precipitation in an attempt to extend the lead-time available. Flood prediction is an important component of flood warning, where the distinction between the two is that the outcome of flood prediction 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 prediction at regional area can be done within seconds by using the technology of artificial neural network. Effective real-time flood prediction models could be useful for early warning and disaster prevention.

1.1 Problem Statement

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.

1.2 Existing System

Physically based models were long used to predict hydrological events, such as storm, rainfall/runoff, shallow water condition, hydraulic models of flow, and further global circulation phenomena, including the coupled effects of atmosphere, ocean, and floods. Although physical models showed great capabilities for predicting a diverse range of flooding scenarios, they often require various types of hydro-geomorphological monitoring datasets, requiring intensive computation, which prohibits short-term prediction. Numerous studies suggest that there is a gap in short-term prediction capability of physical models. For instance, on many occasions, such models failed to predict properly.

1.3 Proposed System

This system aims to collect data from all the states of India and form a generalized dataset. A machine learning algorithm is applied to the labelled dataset, and patterns are extracted, which, in turn, obtain maximum accuracy with real-time input. In general, the dataset collected for predicting is split into a Training set and Test set. Generally, 7:3 ratios are applied to split the Training set and Test set. The Data Model is then created using a Multi-Layer Perceptron Classifier, and the resulting data set is then passed through it for prediction.

1.4 Objective

The objective of Flood Prediction using AI is to design a incremental model to predict floods based on rainfall levels, channel flows, river levels etc. An approach of prediction is using Artificial Neural Networks that has very good working efficiency produces the accurate results. The system helps to improve the performance. This method can be used to predict both long term and short term floods in a given region. Maintaining the project is easy and manageable.

CHAPTER 2

LITERATURE SURVEY

[1] Prediction Analysis of Floods Using Machine Learning Algorithms (NARX & SVM)

The changing patterns and behaviors of river water levels that may lead to flooding are an interesting and practical research area. They are configured to mitigate economic and societal implications brought about by floods. Non-linear (NARX) and Support Vector Machine (SVM) are machine learning algorithms suitable for predicting changes in levels of river water, thus detection of flooding possibilities. The two algorithms employ similar hydrological and flood resource variables such as precipitation amount, river inflow, peak gust, seasonal flow, flood frequency, and other relevant flood prediction variables. In the process of predicting floods, the water level is the most important hydrological research aspect. Prediction using machine-learning algorithms is effective due to its ability to utilize data from various sources and classify and regress it into flood and non-flood classes. This paper gives insight into mechanism of the two algorithms in perspective of flood estimation.

[2] Streamflow Prediction Using Deep Learning Neural Network

The most important motivation for streamflow forecasts is flood prediction and longtime continuous prediction in hydrological research. As for many traditional statistical models, forecasting flood peak discharge is nearly impossible. They can only get acceptable results in normal year. On the other hand, the numerical methods including physics mechanisms and rainfall-atmospherics could provide a better performance when floods coming, but the minima prediction period of them is about one month ahead, which is too short to be used in hydrological application. In this study, a deep neural network was employed to predict the streamflow of the Yangtze River. This method combined the Empirical Mode Decomposition (EMD) algorithm and Encoder Decoder Long Short-Term Memory (En-De-LSTM) architecture. Owing to the hydrological series prediction problem usually contains several different frequency components, which will affect the precision of the longtime prediction. The EMD technique could read and decomposes the original data into several different frequency components. It will help the model to make longtime predictions more efficiently.

[3] Forecasting Daily Precipitation Using Hybrid Model of Wavelet-Artificial Neural Network

Recently artificial neural network (ANN) as a nonlinear interextrapolator is extensively used by hydrologists for precipitation modeling as well as other fields of hydrology. In the present study, wavelet analysis combined with artificial neural network and finally was compared with adaptive neurofuzzy system to predict the precipitation in Iran. For this purpose, the original time series using wavelet theory decomposed to multiple subtime series. Then, these subseries were applied as input data for artificial neural network, to predict daily precipitation, and compared with results of adaptive neurofuzzy system. The results showed that the combination of wavelet models and neural networks has a better performance than adaptive neurofuzzy system, and can be applied to predict both short- and long-term precipitations.

[4] Rainfall Prediction Using Hybrid Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithm

The current rainy season is erratic and very difficult to predict the rain. It requires a method that can predict rainfall with the smallest error as possible. Adaptive Neuro-Fuzzy Inference System (ANFIS) is one of the prediction methods that are quite reliable because it is equipped with a network that can learn. The ANFIS uses Sugeno FIS in its architecture. To improve the prediction results, the Sugeno FIS will be optimized in boundaries of membership function and coefficient consequent rule before it goes into the process of training with ANFIS. A genetic algorithm is used for the optimization process. The results of rainfall prediction using hybrid ANFIS-GA are proven to produce smaller RMSE of rainfall prediction method that has never been done before. With two optimization process in the boundaries of membership function with genetic algorithm and the training process with ANFIS, RMSE values obtained from the rainfall prediction becomes lower. It can be concluded that the results of rainfall prediction using the hybrid method ANFIS-GA produce smaller RMSE compared to the previous methods such as GSTAR-SUR, Tsukamoto FIS, and hybrid Tsukamoto FIS with GA.

[5] Bayesian flood forecasting method

Bayesian forecasting system (BFS) offers an ideal theoretic framework for uncertainty quantification that can be developed for probabilistic flood forecasting via any deterministic hydrologic model. It provides suitable theoretical structure, empirically validated models and reasonable analytic-numerical computation method, and can be developed into various Bayesian forecasting approaches. This paper presents a comprehensive review on Bayesian forecasting approaches applied in flood forecasting from 1999 till now. Results show that the Bayesian flood forecasting approach is an effective and advanced way for flood estimation, it considers all sources of uncertainties and produces a predictive distribution of the river stage, river discharge or runoff, thus gives more accurate and reliable flood forecasts. Some emerging Bayesian forecasting methods were shown to overcome limitations of single model or fixed model weight and effectively reduce predictive uncertainty. In recent years, various Bayesian flood forecasting approaches have been developed and widely applied, but there is still room for improvements.

[6] Comparison of random forests and support vector machine for rainfall forecasting

This study aims to compare two machine learning techniques, random forests (RF) and support vector machine (SVM), for real-time radar-derived rainfall forecasting. The real-time radar-derived rainfall forecasting models use the present grid-based radar-derived rainfall as the output variable and use antecedent grid-based radar-derived rainfall, grid position (longitude and latitude) and elevation as the input variables to forecast 1-h to 3-h ahead rainfalls for all grids in a catchment. Grid-based radar-derived rainfalls of six typhoon events during 2012–2015 in three reservoir catchments are collected for model training and verifying. Two kinds of forecasting models are constructed and compared, which are single-mode forecasting model (SMFM) and multiple-mode forecasting model (MMFM) based on RF and SVM. The SMFM uses the same model for 1-h to 3-h ahead rainfall forecasting; the MMFM uses three different models for 1-h to 3-h ahead forecasting. According to forecasting performances, it reveals that the SMFMs give better performances than MMFMs and both SVM-based and RF-based SMFMs show satisfactory performances for 1-h ahead forecasting. However, for 2- and 3-h ahead forecasting, it is found that the RF-based SMFM underestimates the observed radar-derived rainfalls in most cases and the SVM-based SMFM can give better performances than RF-based SMFM.

CHAPTER 3

SOFTWARE REQUIREMENT ANALYSIS

Software Requirement Analysis in the field of systems engineering and software engineering, encompasses those tasks that are used for determining the needs or conditions to meet for a new or altered product or project, taking account of the possibly conflicting requirements of the various stakeholders, analysing, documenting, validating and managing software or system requirements.

3.1 Feasibility Study

The main objective of the feasibility study is to treat the technical, operational and economic feasibility of developing the application. Feasibility is the determination of whether or not project is worth doing. The process followed in making this determination is called feasibility study. All systems are feasible, given unlimited resources and infinite time. The feasibility study to be conducted for this project involves:

- Technical Feasibility
- Operational Feasibility
- Economic Feasibility

3.1.1 Technical Feasibility

It is the measure of the specific technical solution and the availability of the technical resources and expertise. It is one of the first studies that must be conducted after a project has been identified. A technical study of feasibility is an assessment of the logistical aspects of business operation. This is considered with specifying equipment and software that will successful satisfy the user requirement. The technical needs of the system may vary considerably but should include the facility to produce outputs in a given time, response time under certain conditions and the ability to process a certain amount of transaction at a certain speed.

The proposed system is developed by using Juypter Notebook software. The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more. The "notebook" term can colloquially make reference to many different entities, mainly the Jupyter web application, Jupyter Python web server, or Jupyter document format depending on context. A Jupyter Notebook document is a JSON document, following a versioned schema, containing an ordered list of input/output cells which can contain code, text, mathematics, plots and rich media.

3.1.2 Operational Feasibility

Operational feasibility is mainly concerned with issues like whether the system will be used if it is developed and implemented, whether there will be resistance from the users which will affect the possible application benefits. It is the ability to utilize, support and perform the necessary tasks of a system or program. It includes everyone who creates, operates or uses the system. It is the measure of how well a proposed system solves the problem and takes advantages of the opportunities identified during the scope definition and problem analysis phases. This system helps in many ways. It reduces the burden of maintaining bulk of records of all the rainfall data. Maintenance of the project is also easy and understandable and no major training and new skills are required.

3.1.3 Economic Feasibility

Economic feasibility is the most frequently used method for evaluating the effectiveness of the new system. Economic feasibility is the measure of the cost effectiveness of an information system solution. Without a doubt, this measure is most often and important one of the three. Information systems are often viewed as capital investments for the business, and, as such, should be subjected to the same type of investment analyses as other capital investments.

Economic analysis is used for evaluating the effectiveness of the proposed system. In economic feasibility, the most important is cost-benefit analysis. This project is economical as it mainly depends on the software components which are freely available.

CHAPTER 4

SYSTEM REQUIREMENT SPECIFICATION

The main purpose of System Requirement Specification is to translate the ideas in the minds of a client into a formal document. Through System Requirement Specification the client clearly describes what it expects from the proposed system and the developer clearly understands what capabilities are required to build the system. It includes a variety of elements (see below) that attempts to define the intended functionality required by the customer to satisfy their different users. The purpose of this document is to serve as a guide to developers and testers who are responsible for the development of the system.

4.1 Functional Overview

- Rainfall data is collected from meteorological websites.
- An algorithm is used for creating a model.
- The chances of a flood are predicted.

4.2 Operating Environment

Operating environment involves minimum software and hardware requirements required by the system.

4.2.1 Software Requirements

- Operating System : Windows 7 or above.
- Tools used : Jupyter 6.0 or above.
- Programming Language : Python Programming Language.

4.2.2 Hardware Requirements

- Processor : Pentium 4 or above
- RAM : 6GB or above
- Hard Disk : 10GB or above
- Input device : Standard Keyboard and Mouse.
- Output device : High Resolution Monitor

4.3 Functional Requirements

Functional Requirements defines the function of a system or its component. A function is described as a set of inputs, the behavior and outputs. Functional requirements specify particular results of a system. Functional requirements drive the application architecture of a system. Following are the functional requirements used in the project.

- Process data in any required format i.e., xls,csv etc.
- Import the data and store its metadata relating to the segmentation without corrupting the contents.
- Concepts and target extraction.
- Predicting the outcome.

4.4 Non-Functional Requirements

Non-Functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. Non-Functional requirements are often called as quality attributes of a system. The following are the nonfunctional requirements of the application.

- Availability: The System will work as required according to the specified requirement.
- Reliability: The System has the ability to consistently perform the intended or required functions.
- Maintainability: The maintenance of a functional unit can be performed in accordance with the prescribed requirements
- Accessibility: The System can be accessed by any appropriate users.
- Operability: The System has the ability to keep itself safe and reliable, according to predefined operational requirements.
- Usability: The System is ease of use and learnable by the users.
- Responsiveness: The system has ability to respond to the user very fast as soon as the input is fed.
- Cost: The system is economically feasible.

4.5 Performance Requirements

This application system will avoid use of papers and going through each and every detail for each state or district. Thus, this application results in efficient usage of time. This application will be useful for individuals, city planners, researchers etc. Application calculates the result automatically and displays it to the user. Interface is designed in such a way that user can understand very easily. Maintenance of the project is easy and understandable.