

---

# Flood Detection And Notification System Based On Decision Tree Algorithm

**Abstract**—Floods are common catastrophic events that cause extreme devastation of any nation. They are usually brought about by precipitation and overflow of waterways, particularly during periods of excessively heavy rainy season. This project aims to monitor the flood condition and send alert if there is an occurrence of risk through IoT. The measurement of rising water level is done to detect the flood. The system uses three sensors to detect temperature, humidity and water levels at every stage. The detected sensor values are transmitted to IoT through Wi-Fi from PIC microcontroller. The system instantaneously uploads and broadcast sensor values through cloud. The decision tree algorithm is implemented to perform the classification process. The experimental results for Correctly Classified Instances and Minimum Absolute Error indicates that the proposed Decision Tree Algorithm gives 99.6% classification accuracy with Minimum Mean Square Error than HyperPipes Algorithm.

**Keywords**—Flood Detection System, Sensor Data, PIC Microcontroller, IoT, Decision Tree Algorithm.

## Introduction

Flood is considered as one of the most common natural phenomenon in the world, impacting human lives and making extreme damage to goods and properties. Avoiding floods before severe damage can provide sufficient time for inhabitants evacuate in the nearby areas. The Prediction and warning system could have potential impact to reduce the severe of flood affects.

This natural disaster can't be gotten away however appropriate managing and pre-alarming system can reduce its seriousness. In most of the developing countries, flood monitoring cell are not properly equipped with smart and adaptable flood alarming system. As a result, people from flood affected areas are suffering the consequences of flood every year. The main objective of this project is to develop and design a flood detection system that will detect flood automatically and transmits data through IoT.

This system is used to detect the current water levels on flood by taking sensor values from outside environment and it will give real-time information to the appropriate station about severity. At the second stage, an experiment is made based on machine learning algorithm to analyse the level of flood data to quantify if the level of water is normal or dangerous condition. Several values are collected from sensor data which is trained and classified using Decision Tree algorithm. Performance metrics such as correctly classified instances, incorrectly classified instances, Kappa statistics, Mean Absolute Error, Relative mean absolute error are obtained for several sensor data values shows better classification accuracy than the Random Forest algorithm.

## Literature Review

Do et al have proposed an early flood warning framework has been examined and developed.

---

The framework comprises of an early flood cautioning station which is charged by solar power. The flood cautioning station is outfitted with the water level sensor and the precipitation sensor. The information, which is gotten from sensors, will be put away in memory card and transmitted to observing focus by means of GPRS convention. The monitoring center's function is to collect, investigate information from the stations and send results to notification center where cautioning can be produced to the thought about local locations. The warning levels of flood can be set up easily by utilizing the SMS message from client's phone. A webpage was made to show the water level, precipitation and the anticipated data for continuous action. The problem with this system is setting up the conditions for testing in the laboratory takes several months to have the test results in reality[1].

Sakib et al have proposed a neuro-fluffy controller based flood observing framework utilizing remote sensor organize. The framework utilizes IEEE 802.15.4 convention as a conveyed hub to gather the sensor data, for example, water level information from the waterway, precipitation, and wind speed information from a chosen site. So as to approve the flood checking framework, Chadpur, a flood inclined location of Bangladesh, has been considered as chosen site. The sensors data are sent to the distributed alert center through Arduino microcontroller and the Zigbee Transceivers. Sensor data are investigated by the neuro-fuzzy controller utilized in Raspberry Pi microcomputer to report the flood alarms. The remote sensor arrange is associated as work topology which can send motions over far separation. It is incompetent to control moving procedure with time delays and change in the MF can require a change in the rules. And it undergoes multi-parameter optimization problem[2].

Mousa et al have designed new sensing device that can continuously screen urban flash floods and traffic blockage for flood observing. This detecting device depends on the blend of ultrasonic range finding with remote temperature detecting, utilizing a mix of L1-regularized regression and artificial neural networks to process estimation information. Second, comparing calculations have been implemented on a low-control remote sensor platform. The outcomes show that urban water levels can be dependably evaluated with reduced error under 2 cm, and that the pre-processing and AI plans can keep running progressively on as of now accessible remote sensor stages. In the present case, ANNs show great combination properties however it empowers us to utilize a low number of neurons and layers, making it appropriate to a low power embedded systems application[3].

Khalaf et al have proposed a depiction of an alert creating framework for flood detection with an emphasis on deciding current water level utilizing sensors. The system at that point gives warning message about water level by means of Global Communication and Mobile System modem to specific person. Other than the Short Message Service, the system quickly transfers and communicate data through web base open system. AI calculations were directed to play out the classification process. The tests were done to characterize flood information from typical and in danger condition in which arrangement precision was accomplished utilizing Random Forest algorithm. The calculation experiences complexity and tedious to develop than decision trees[4].

Dashpute et al have proposed a flood detection system to screen rising water in local locations. Utilizing ultrasonic sensor they made flood level detecting device which is connected to Node MCU controller to process the sensor's analog signal into a usable advanced estimation of separation. The client can get real-time data on observing overflowed streets over SMS based administration. The level of the flood will be partitioned into four. The flood sensor and

---

microcontroller will be controlled by a solar based energy to assist constant activity of water flood height identification and system information transmission. The Arduino Flood Detector System is created to screen flood that will support drivers or street client to stay away from issue when flood happened. The issue distinguished in this system is, it will not predict the values or it takes previous for classification by using any machine learning algorithm[5].

## System Design & Description

The above mentioned papers reveal about the concepts of flood detection in different methodologies. There are several equipment that can be combined together to send notification using water level sensor to the emergency management authority. In all the above papers, they defined different approaches only for knowing current statistics of sensor values for alert which does not concentrate on early prediction by taking several data's utilizing algorithms for flood detection.

The project proposes the design and development of flood detection and notification system that will detect the flood automatically using sensors. To detect the current level of the flood where the system sensor will be divided into different levels at every stage. The collected values are transmitted to cloud through Wi-Fi module from PIC microcontroller for data interpretation.

The proposed system uses decision tree algorithm for classification process and to analyse the level of flood data to notify if the level of water is normal or risk condition.

### PIC Microcontroller

DHT11

WIFI Module

Water level 1

Water level 2

IOT

Water level 3

Fig.1 Proposed Block Diagram

Fig. 1 shows the block diagram of flood detection system. PIC16F877a microcontroller is the main controlling unit in this system. The three water level sensors are placed at different levels to determine the current water level using predefined values. The water level 1 detects the normal level of water, while the water level 2 detects the above water level, and the water level 3 detects the risk condition of water level. Also DHT11 sensor is connected with microcontroller to detect the temperature and humidity of current location. Data collected from sensors are transmitted to the main control unit. PIC receives various signals from sensors and transmits those sensor values to the cloud through Wi-Fi module. The framework continuously uploads the sensor data to the cloud[6].

---

## **PIC16F877a Microcontroller**

The PIC16F877a finds its applications in a huge number of devices. It has total number of 40 pins and there are 33 pins for input and output. The PIC16F877A features with 256 bytes of EEPROM data memory, an ICD, Self-programming. It has two comparators, 8 channels of 10-bit Analog-to-Digital converter (ADC), 2 capture/compare/PWM functions, the synchronous serial port can be configured as either 3-wire Serial Peripheral Interface or the 2-wire Inter-Integrated Circuit bus and a Universal Asynchronous Receiver Transmitter (USART).

In this project, the PIC16F877a microcontroller is the main controlling unit. The flood detection and warning is done based on code programmed from PIC microcontroller. To measure temperature and humidity, the DHT11 sensor is connected with microcontroller for finding the current climatic condition. The probe sensors are interfaced with microcontroller to measure current water level at each stage. The detected sensor data are transmitted to cloud using ESP8266 Wi-Fi module through PIC microcontroller.

## **Three level Probe Sensor**

The Probe Sensor works on the principle of conductive level sensor in which the resistance between two measuring electrodes can be change when there is conductivity between water and the ground. When the water covers the electrodes, it forms an electric current, causing current to flow. In this project, three level Probe sensor is used to measure water level at three stage. The probe sensor 1 detects the normal water level, while the second probe sensor detects above water level and the third probe sensor detects the risk condition of water level. These are interfaced with PIC microcontroller to process the obtained sensor values.

## **ESP8266 Wi-Fi Module**

In this project, the ESP8266 Wi-Fi module is used to provide network between controller and the cloud. The detected sensor values are transmitted to the cloud through Wi-Fi module.

## **Proposed Methodology**

There are two main parts in experimental methodology. The initial part of the experiment is hardware setup for collecting flood data from different sensors and it is uploaded in the cloud periodically. The second part of the experiment is completely based on machine learning approach for analyzing the level of data to quantify the flood level in order to detect whether the level of water is normal or dangerous condition[7].

- Start
- Establishing Connection between Controller and Network
- Read Sensor values
- Display on LCD
- Upload the values on Cloud
- Decision Tree Classification
- Stop

Fig.2 Flow Chart for Proposed Methodology

---

In order to collect datasets for classifying flood data with normal and at risk condition, raw data has been extracted from sensor values for around 100 flood values initially. Further, it could be classified by means of Decision Tree Algorithm to quantify the results obtained.

## Decision Tree Algorithm

A decision tree strategy pursue a tree like structure prescient model, where each interior hub speaks to a property, each branch speaks to a consequence of the test, and class label is represented by each leaf hub or an end hub. It can take care of the issue of both classification and regression. In a decision tree, each interior hub parts the occurrence space into at least two sub-spaces as indicated by a specific discrete function of the attribute characteristics esteems. In the least complex and most regular case, each test thinks about a solitary trait, with the end goal that the case space is divided by the characteristic's esteem. On account of numeric characteristics, the condition alludes to a range. It is anything but difficult to change over decision trees into characterization rules. Decision tree learning uses a decision tree as a predictive model which maps perceptions around a thing to decisions about the thing's objective esteem. Decision tree classifiers get comparative or better precision when contrasted and other arrangement techniques. Finding the best attribute is based on two specifications mentioned below. In order to choose the best attribute, the value among the dataset should have high value of information gain.

## Entropy

Entropy is the proportion of homogeneity in the information. Its esteem is ranges from 0 to 1. Its esteem is near 0 if all the precedent has a place with same class and is near 1 is there is practically equivalent part of the information into various classes. Presently the equation to ascertain entropy is:

$$\text{Entropy}(S) = (1)$$

Here  $c$  corresponds to different types of classification and  $N_i$  represents the proportion of the data with  $i$ th classification.

## Information Gain

Information Gain measure the decrease in entropy by arranging the information on a specific attribute. The formula to calculate gain by splitting the information on Dataset's' and on the attribute 'a'.

$$\text{Gain}(s, a) = (2)$$

Where  $\text{Entropy}(S)$  represents the entropy of the dataset and the  $S_v$  denotes the weighted entropy of different possible classes. The attribute which contain high information gain among the dataset is selected as an internal node for tree doing further classification.

## WEKA

---

WEKA (Waikato Environment for Knowledge Analysis) workbench is set of different data mining tools developed for machine learning process. For simple access to this usefulness, it contains a collection of visualization tools and algorithms for data analysis and predictive modeling together with graphical user interfaces.

## **Experimental Results**

### **Experimental Setup**

Fig. 3 shows the experimental setup of this system. The implementation of flood detection and notification system has been proposed. The main controlling unit PIC microcontroller and different sensors are been connected with controller to detect the sensor data values. Finally, these values are been send to cloud through ESP8266 module.

Fig.3 Experimental Setup

### **Hardware Results**

The detected sensor values are displayed on LCD which is shown in Fig. 4.

Fig.4 Sensor values display on LCD

Also these data are transmitted to IOT through Wi-Fi module. The system continuously uploads and broadcast sensor values to the cloud about water condition.

[image: C:\Users\VINOTHINI\Desktop\phase2 papercloud 1.png]

Fig.5 Values in cloud for temperature and humidity

Fig. 5 shows the detected temperature and humidity sensor values uploaded on the cloud

[image: C:\Users\VINOTHINI\Desktop\phase2 papercloud 2.png]

Fig.6 Values in cloud for three level Probe sensor

The detected three level Probe sensor values are uploaded on the cloud is shown in Fig. 6 which shows the normal or risk condition of flood level.

### **Classification Results**

Around 100 Datasets of different sensor values are been taken and it is processed in this proposed work. The Decision Tree Machine learning algorithm is implemented for classifying flood data efficiently by utilizing Weka Explorer platform. Various parameters are been measured and tabulated for obtained sensor data.

### **Confusion Matrix Analysis**

To examine the proposed machine learning method for classification of flood data, a confusion

---

matrix has been evaluated to analyse the dataset taken from flood data. In this analysis, the two algorithms were implemented such as decision tree and HyperPipes algorithm.

The True Positive measures the correctly classified instances (or predict the positive class) while, True Negative illustrate the negative class instance respectively. False Positives gives negative symbols while, False Negatives illustrates the positive instance. The term Accuracy defines the correctness or true value of positive instances and it is calculated as follows:

(3) While equation (4) is used to calculate the proportion of positive instances to find correctly classified instances.

(4) Equation (5) is used to calculate the proportion of negatives that were incorrectly classified instances.

(5) Equation (6) shows true negative to calculate correctly classified positive instances.

(6) Equation (7) shows False negative to calculate incorrectly classified negative instances.

(7)

## **Comparative Analysis of Classifier Output**

Parameters

HyperPipes Algorithm

(Existing Algorithm)

Decision Tree Algorithm

(Proposed Algorithm)

Correctly Classified Instances

89.35%

99.62%

Incorrectly Classified Instances

10.65%

0.38%

Kappa

statistic

---

0.7942

0.9835

Mean absolute error

0.4599

0.0212

Root mean squared error

0.4667

0.0910

Relative absolute error

93.6543%

3.9426%

Root relative squared error

92.6775%

19.8559%

Total Number of Instances

805

805

Table 1 shows the comparative analysis of proposed Decision Tree Algorithm which processed using Weka Explorer.

From the above results, it is inferred that the Decision Tree Algorithm implemented in machine learning software produces correctly classified instances with 99.6% of accuracy while, the incorrectly classified instances are 0.38% respectively with minimum Mean absolute error than the HyperPipes Algorithm. The algorithm produces high classification accuracy during the prediction of tree attribute because of calculating the information gain and entropy efficiently even for massive datasets. Its specificity can be increased by validating all possible outcomes of each decision class label effectively. Also it reduces overfitting problem and easy to set parameters during classification.

The accuracy can be defined by correctly classified instances or occurrence of true value outcomes after classification. Further, it could be determined by calculating True Positive, False Positive, True Negative and False Negative respectively.



---

## Accuracy of Classification

### Accuracy By Class

#### HyperPipes Algorithm

(Existing Algorithm)

#### Decision Tree Algorithm

(Proposed Algorithm)

#### True Positive

1

0.993

#### False Positive

0.165

0.013

#### Precision

0.125

0.998

#### Recall

1

0.993

#### ROC Area

0.982

0.997

Table 2 shows the accuracy of classification by class in terms of parameters such as True Positive indicates sensitivity while, False Positive indicate the specificity and further Precision, Recall, ROC area has been calculated.

From the above results, it is inferred that the proposed Decision Tree Algorithm produces very low true positive and false positive rate, hence the classification accuracy is very much greater than the HyperPipes algorithm. In order to reduce the true positive rate, the classifier statistically

---

derive the outcome where the model correctly predicts the positive class or correctly classified instances while for reducing false positive rate, the classifier should derives the negative class efficiently.

## Conclusion

The design and implementation of a flood detection and notification system based on Decision Tree algorithm is proposed in this work. The system divides the flood level into three stages depending on the measured data collected from sensors. For each warning level, the system continuously updates and broadcast the sensor data to the cloud from PIC microcontroller. The implementation and comparison of Decision Tree and HyperPipes Algorithm are made for around 100 datasets of sensor data collected from different sensors. The experiments were carried out to classify flood data from normal and at risk condition in which the proposed Decision Tree algorithm gives better classification accuracy with lesser possibility of error than HyperPipes classification.

## References

1. H. N. Do, M. Vo, V. Tran, P. V. Tan, and C. V. Trinh, "An Early Flood Detection System Using Mobile Networks," pp. 599–603, 2015.
2. S. N. Sakib, T. Ane, N. Matin, and M. S. Kaiser, "An Intelligent Flood Monitoring System for Bangladesh Using Wireless Sensor Network," pp. 979–984, 2016.
3. M. Mousa, X. Zhang, and C. Claudel, "Flash Flood Detection in Urban Cities Using Ultrasonic and Infrared Sensors," vol. 16, no. 19, pp. 7204–7216, 2016.
4. M. Khalaf, A. J. Hussain, D. Al-jumeily, P. Fergus, and I. O. Idowu, "Advance Flood Detection and Notification System based on Sensor Technology and Machine Learning Algorithm," pp. 105–108, 2015.
5. K. R. Dashpute, V. B. Gaikwad, and S. S. Sawkar, "Flood detection using iot," IJARIIIE-ISSN(O)-2395-4396, no. 2, pp. 1289–1292, 2018.
6. B. K. Durga, A. Ece, S. Peter, and S. E. College, "Design of Early Warning Flood Detection Systems," vol. 5, no. 4, pp. 794–799, 2018.
7. E. Basha and D. Rus, "Design of Early Warning Flood Detection Systems for Developing Countries," pp. 1–10, 2004.