FLOOD FORECASTING SYSTEM USING DEEP NEURAL NETWORK

Chandini Unnikrishnan[1], Rakshitha S[2]

[1]Professor, Dept. of Computer Science and Engineering, Atria Institute of Technology, Karnataka, India
[2]UG Student, Dept. of Computer Science and Engineering, Atria Institute of Technology, Karnataka, India

---***---

Abstract - *Basically flood is both natural and ecological process which occurs when water inundates land, overflow of water, gathering of water in the land basins and also due to drainage. This influences damaging of human, flora and fauna's lives, environmental destruction, mental depressions on human and also economy of the country. Today India is also one of the flood affected nation among the world with the best example of recent flood in Kerala on August 2018. There are wide technologies such as Internet of things (IOT), Machine Learning (ML), used in prediction of flood based on the parameters such as temperature, rainfall, water level, humidity, pressure, etc. But the real challenge is to estimate the occurrence of flood with the limited number of parameters such as rainfall intensity and air intensity and also alert people. Therefore according to Deep Neural Network (DNN), flood is forecasted based on only few parameters. In addition DNN is also compared with K-Nearest Neighbor (KNN) in terms of accuracy, precision and recall. The overall outcome taken from this is to effectively forecast the flood based on the calculations on the intensities only before the occurrence of flood and also to alert people through online flood portal system so that many dangerous effects could be stopped.*

***Key Words*: Deep Neural Network, K – Nearest neighbor, accuracy, precision, recall.**

1. INTRODUCTION

Flood is one of the dangerous environmental effects which occur due to many reasons such as, when there is a high volume of water carried by rivers, creeks and many geographical features to the areas where water cannot be drained easily, accumulation of water in the river basins where when there is a heavy rainfall which again results in flood, drainage system are not consistently checked by civil department severely hinders the main operation of acceptable drainage system. As India is one of the second highest populated nations, flood impacts high destruction by causing more fatalities on lives and properties. This can also results in completely destroying under developed areas. According to the calculation of Indian geographical areas, around 3295 lakh hectares of lands are inclined to floods. This results on a high amount of risk. In recent years, there are wide technologies in Information Technology (IT) easing in flood management. Example, there are many parameters are taken into considerations and using respective sensors in manipulating each process. The same goes in practicing of geological and meteorological data related to flood, where Artificial Neural Network (ANN) comes into picture in prediction. The best well–known model used in ANN in analyzing flood is by using a technique called Adaptive Neuro–Fuzzy Inference System (ANFIS) model where, this warns people or respective station in flood occurrence and also flood possibility can be analyzed or determined. With a best emergence of IOT and ML algorithms, various works and experimentations are carried out in implementing sensors and machine learning algorithms, such as ANN for finding out flood with an alert system on go. Various works have been in progress on the use of ANN algorithms for forecasting flood through the parameters such as water level, flow, humidity, waves speed, dryness of earth, and so on. But the bigger disadvantage is that in all those systems the flood monitoring systems have focused on ANN by using an approach of single hidden layer on the above parameters. Moreover, ML algorithms such as linear regression, SVM, Forest classifiers are also employed accordingly. However, none of the proposed models planned on the parameters such as rainfall and air intensity for predicting in an early stage. More concentrated parameters are water level which will not be constant and not sufficient in predicting flood. flood early warning for India where only air and rainfall intensities taken into consideration. This is easily compared with one of the best ML algorithm called KNN by means of accuracy, precision and error at predicting flood at an early stage before occurrence. The main aim of the paper is to collection the dataset pertaining to air intensity, rainfall intensity and merging them into a single dataset, validation of KNN, validation of DNN, comparative analysis of KNN and DNN for accuracy on the flood prediction, and also a warning alerts as danger to respected people and departments.

2. LITERATURE SURVEY

Various parameters of basic physiological parameters of nature are considered in forecasting of flood.

P. Mitra et al (2016) investigated forecasts flood by using both IOT and ANN in such a way that, it uses many sensors for checking many parameters and also uses zig-bee configuration with a micro controller unit. Rainfall, water level, humidity, pressure parameters are collected from WSN model and a log sigmoid curve is plotted using MATLAB Neural Network Tool box software. Based on the threshold value coefficient correlation on each process is validated which is a long and a tedious process.

S. Bande and V. V. Shete (2017) considered Raspberry module acts as IOT node with Wi-Fi using WLAN technology in a Wi–Fi hotspot and may obtain

internet access using it. ANN is used for analysis of data with NARAX architecture with step consideration alert.

S. I. Abdullahi, M. H. Habaebi and N. A. Malik (2018) forecasting is done with IOT where a Node MCU ESP8266 enables transmission of data collected from sensor on the Thing Speak Channel for best visualization. And when it comes to ANN Microsoft's Azure Machine Learning is used to predict flood occurrence. And the main result shows percent of accuracy and precision.

3. PROPOSED SYSTEM

Here the datasets of rainfall and air intensities are collected on the daily aspects of weather checking. There are different classification of both air and rain intensities which are also collected. Based on these parameters the datasets are merged by classifying training and test splits. These are also pre-processed. Pre-processing takes place by computation with KNN and DNN. The percent of results of two techniques are compared one another through accuracy, precision, and recall calculations based on the confusion matrix. By analyzing those value bar plots are also plotted graphically where an alert can be verified. These alerts of DANGER predicts flood through online flood portal system by Meteorological departments. The innovative thing here is to collect few parameters instead of validating various parameters which keeps changing constantly without any guarantees. Here the main ability to perform all these tasks is done by DNN where it can collect unstructured data which overkills complexities with an activation function Leaky ReLU which results effectively even considering vast amount of data. These signals for flood prediction level through different weights which can save many lives over deterioration.
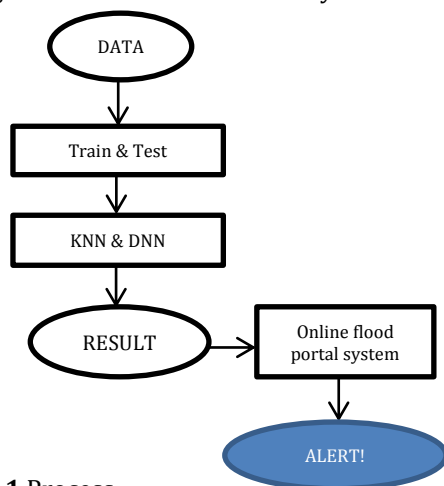

Fig -1 Process.

4. METHODOLOGY

The flood surveillance and an alert system are done with an advance supervised Machine learning algorithm called KNN and also compared with DNN with analysis

of the flood warning. When the alerts are sent, where the information is shared to those people by departments of meteorology by browsing the effect of overall percent of accuracy with compared with both DNN and KNN.

A. KNN

KNN is one of the simplest classification algorithm which is easier to pick up. KNN can also be used for classification purpose where the output is a class membership. This mainly use the datasets defined in which dataset are separated into several classes to predict the classification of a new sample point. Here, this is mainly done on feature similarity between the two different data as KNN works effectively on the feature similarity also. Steps are as follows, Steps to implement KNN model:

1. Load the dataset collected with the real evidence.
2. Determine the value of k. To determine this there are certain procedure that is,
   - If the k initialized with lower value, more noise will be resulted in the output which results in over fitting.
   - If the greater value is chosen the computational effect and logic on KNN will sometimes lost.
   - Therefore the perfect value can be can be chosen by using the formula
     $$k = \sqrt{n}$$
3. Calculate Euclidian distance where, the distance between each test data and training data is noted by using $\sqrt{(q_1^2}$
   $$d(p,q) = d(q,p)$$
4. $= \sqrt{((q_1 - p_1)^2 + (q_2 - p_1)^2 + \text{------}(q_n - p_n)^2)}$

   $= \sqrt{\sum_{i=1}^{n}(q_i - p_i)2}$

5. Arrangement of values according to ascending order
6. Obtain the nearest class of each row for prediction.
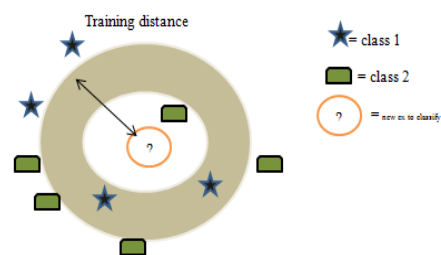7. The best predicted class is returned in figure [2].

Fig -2 KNN classification.

B. DNN

Perceptron in deep learning:
Basically perceptron proposed by Frank Rosenblatt in 1957. It is defined as neural network unit which performs computational tasks to detect features and business intelligence in the input data. Perceptron more than three layers constitute deep learning.

Basically perceptron is the simple model of neuron. Single Layer Perceptron is where the computation of a single layer is performed over the calculation of sum of the input vector as in (4) each with the value multiplied by corresponding element of vector weights. The value which is displayed in the output figure [3] will be input of an activation function.
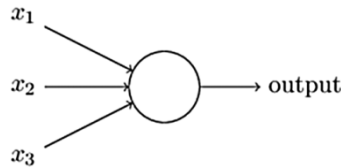

Fig -3 Single layer Perceptron

Multi-layer Neural Network (MLP)
This is basically defined as a class of feed forward neural network. This is also sometimes used many times in artificial neural network. MLP with highly connected are referred to threshold activation functions. This mainly consists of three layers such as, input layer, hidden layer and output layer. Each node of input layer uses non-linear activation function. MLP is differentiated from Single layer perceptron through non-linear activation functions with two output as shown in figure [4].
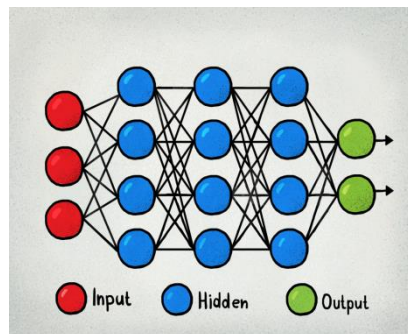

Fig -4 DNN structure with output

$$weight = weight + learning_{rate} * (expected - predicted) * x$$

Here, MLP is used because, each unit of its perceptron receives multiple input signals and threshold value is set. If the value of the input signal exceeds the set threshold value, it will output a signal or sometimes do not proceed with the output. As DNN and KNN are supervised here, perceptrons plays a major role on supervised algorithms and classification. This can be then used to predict the class of sample. The major advantage here is to classify the given input parameters. Now these, MLP is proceeded with an activation functions which are really important to learn and make a sense of something which are more complicated and Non-linear complex functional mappings between inputs and response variable. They also connect non- linear properties in our dataset.

Model works according to below steps:
1. Data preparation: dataset collected is used to demonstrate the implementations of DNN. Nominal attributes should be converted into binomial attributes.
2. Modeling of parameters: Training dataset is connected to neural net operator which accepts real data types and normalizes the value.

   - Hidden layer: determines the number of layers, size, name for easy identification in the result. Default size is -1 which can also be written by specifying a number.
   - Training cycles: this is where a number of times training cycles repeat. Default is 500.
   - Learning rate: this takes and determines value from 0 to 1. 1 represents error and 0 represents new weight added to previous weight and has less error correction.
   - Momentum: is used to obtain globally optimized results by adding a fraction of old weight to the new weight.
   - Decay: this reduces the value of the learning rate and brings it closer to 0 for the last training record.
3. Normalization
4. Evaluation
5. Execution and interpretation: This shows output topology. Percent of training accuracy is yielded. The one which are incorrectly classified are also shown easily.

Syncing with Activation Function called Leaky ReLU
   - The various activation functions in neural networks are Linear Activation function, non-linear, sigmoid, Tanh, ReLU, Leaky ReLU. Among those, here Leaky ReLU is chosen because, it fixes the "dying ReLU" as it doesn't have 0 slope parts. And also it speeds up training, where a mean activation closer to 0 makes training faster.
   - Customized version of P ReLU is leaky ReLU. Constant multiplier α is equal to 0.1 for this customized function. This sometimes varies to 0.01 in []. In each session alpha value is picked for the calculation.
   - Leaky ReLU is defined to address this problem. In ReLU function x is defined to be 0 for negative values of x, we define it as an extremely small component of x, the mathematical expression is :

$$f(x) = 0.01, \qquad x < 0$$
$$= x, \qquad x \geq 0$$

Based on the figure [5] function works as, the value x if lesser than 0 output multiplied with 0.01 or it returns value normally.

Comparison of KNN with DNN through precision, accuracy, recall done through confusion matrix in table [1]. The formulas are shown below,

$$PRECISION = TP/(TP + FP)$$

$$ACCURACY = (TP + TN)/(TP + FP + FN + TN)$$
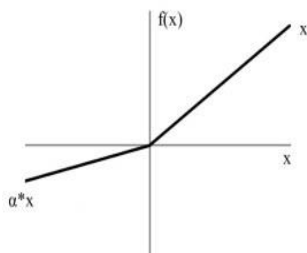
$$RECALL = TP/(TP + FN)$$

Fig -5 Leaky ReLU Plot

ARCHITECTURE OF AN EARLY FLOOD WARNING SYSTEM USING DNN AND KNN SYSTEM:

Major step is to collect the datasets pertaining towards those areas nearer to rivers or oceans. When these datasets are collected, construct a model fig [7] which has to be split into training and test data and apply suitable KNN and DNN model for the analysis of the result by means of accurate percentage and precision calculations. If a high bias is seen it must be redefined to corresponding steps, until a lower bias is seen. Accuracy obtained after biasing leads to setting up of model. And the secondary part is online flood portal system where an admin with meteorological team with respected authorities are connected as shown in figure [6]. Regular updates are seen on every action before flood occurrence.
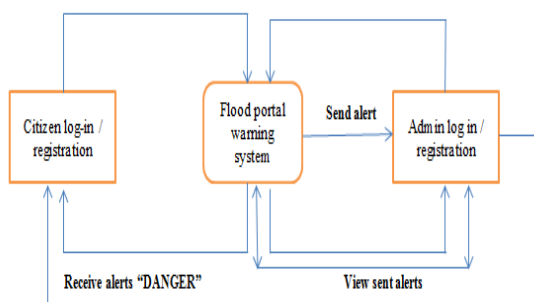
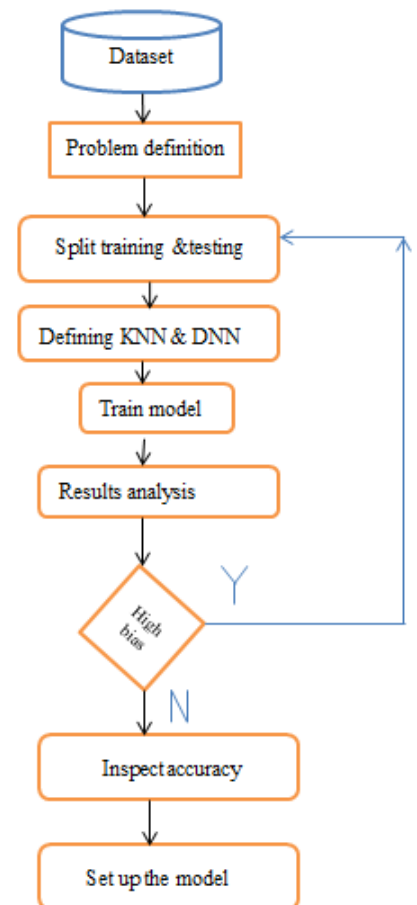**Fig -6** Flood portal system

Fig -7 System architecture.

5. RESULT AND DISCUSSION

Checking with confusion matrix: It is a table that is often used to summarize the performance of classification model on the set of test data with the expected outcome values. To construct a set of predictions on each row is made. From those expected outcomes and prediction count, correct positive predictions on each class is obtained. At the end the numbers of possible correct and incorrect predictions are summarized with count values based on table [1]. TP and TN denote the number of instances which have been correctly classified as no flood occurrence and flood occurrence respectively. FP and FN signify the number of instances which have been wrongly classified as no flood occurrence and flood occurrence respectively.

The prediction of accuracy and other related metrics for each algorithm are given in Table [1] and Table [2] and the bar graph figure [8] is plotted which consists of KNN and DNN verification, so this helps to predict the flood easily, so that necessary precautions are taken in order to save many lives. The accuracy of prediction will be pre-dominantly high for DNN compared to KNN or any other ML algorithms. This indicated DNN is one of the suitable technologies in

prediction of flood or forecasting it with only few parameters called air and rain intensity. Based on the accuracy, precision, recall flood can be easily predicted with a few easy steps and can be alerted danger.

Y = 0	Y = 1		
TRUE NEGATIVE	FALSE POSITIVE	X = 0	TRUE
FALSE NEGATVE	TRUE POSITIVE	X = 1	LABEL

TABLE [1] Confusion Matrix

MODEL	ACCURACY	PRECISION	RECALL
KNN	85.73%	91%	90%
DNN	92%	96%	94%

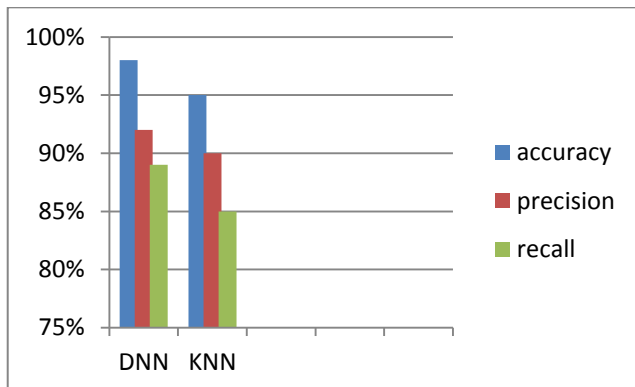TABLE [2] Comparison of KNN and DNN model.

Chart -1:   Comparison graph

6. CONCLUSION AND FUTURE ENHANCEMENT

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.

7. ACKNOWLEDGEMENT

8. REFERENCE

[1] P. Mitra et al., "Flood forecasting using Internet of things and artificial neural networks," 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, 2016, pp. 1-5.

[2] S. Bande and V. V. Shete, "Smart flood disaster prediction system using IoT & neural networks," 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bangalore, 2017, pp. 189-194.

[3] S. I. Abdullahi, M. H. Habaebi and N. A. Malik, "Flood Disaster Warning System on the go," 2018 7th International Conference on Computer and Communication Engineering (ICCCE), Kuala Lumpur, 2018, pp. 258-263.