International Journal of Advanced Science and Technology
Vol. 29, No. 6, (2020), pp. 2804 - 2812

Prediction of Flood by Rainfall using MLP Classifier of Neural Network Model

Dr. V.V. Ramalingam*[1], Rohan Mishra[2], Sagar Parashari[3]

[1]*Faculty, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India*
[2,3]*Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India*
[1]*ramalinv@srmist.edu.in**, [2]*rohanmishra0010@gmail.com*,
[3]*sagarparashari198@gmail.com*

Abstract

Floods are one of the most cataclysmic events in the face of Earth, which are exceptionally unpredictable to demonstrate and difficult to predict at an earlier point of time. The examination of improvement of flood-forecast designs has added to diminish hazard, policy suggestion, minimization of the loss of lives and damage caused to properties pertaining to floods. To reenact the compound numerical interpretations of physical actions of floods, during the last two decades, neural system blueprints have assisted in improving and advancing flood prediction structure providing better execution and cost-effective solutions. This project helps predict the occurrence of a flood by rainfall dataset with neural network-based techniques to prevent this problem of floods. The analysis of the dataset by Multi-Layer Perceptron Classifier (MLP) is carried out to acquire certain information like; variable identification, treatment of missing values, data validation and data cleansing and preparation will be done on the complete provided dataset. The performance of the algorithms utilized for flood forecasting is seen by the accuracy computation with an evaluation classification report, confusion matrix identification and result. It shows the effectiveness of the graphical user interface-based application by predefined attributes and gives an early alarm for an impending disaster.

Keywords: *SK-learn, rainfall, disaster, natural language processing, forecast.*

1. Introduction

The objective is to build an AI model for ongoing flood estimation, to almost certainly supplant the versatile multi-layer perceptron (MLP), neural system models, by foreseeing brings about the type of exactness. In urban communities, sub-passes, or low-lying areas are generally helpless against water logging and waste issues. In such flood-vulnerable points, water gets gathered pointlessly in a brief span of time. Surface spillover, relative rise and a deficient section of water to sewerage are important focuses on the situation of water logging. In this way, estimating floods at such locations is significant and dire. The run of the mill techniques of flood gauging is upscale and profoundly intricate. Prediction of weather, along with rainfall prediction, is a significant task behind the forecast of a flood. Climate determining includes re-enactments dependent on material-science and differential conditions. The forecasting of rainfall is done using satellite imaging radars. A Doppler climate radar is utilized to find the amount of rainfall received and recognize the movement of rain droplets. A committed climate satellite gives pictures utilizing which data about precipitation can be reasoned and it's for momentary flash-flood forecast in urban regions is to set up a model assimilating the factors affecting flood and utilize the intensity of AI systems to evaluate flood early. Fast urbanization, environmental development, and intense precipitation have brought about developing instances of flash-

floods in metropolitan areas. It is imperative to foresee the event of a downpour so that its fallout is limited. According to its name, an urban flash-flood happens in a metropolitan territory in an extremely limited period of time. To diminish the effect of these occasions, transient determining or now-casting is utilized for forecasting of the very near upcoming occurrences. In customary strategies for flood gauging, current climate conditions are inspected utilizing regular techniques, for example, the utilization of radar, satellite imaging and several evaluations including confounded scientific conditions. Be that as it may, late improvements in Machine Learning(ML) and Information and Communication Technology (ICT) have aided us in contemplating this issue from an alternate point of view. The fact of the matter is to structure a model keeping in mind the criterion causing the above type of flood and foresee the same ahead of time.

2. State of the art

Various procedures have been actualized in the field of flood forecast, which incorporates different AI calculations. While trying to improve the exactness and accelerate the process of flood prediction to prevent loss of human life and property, many methodologies have come up based on the different arenas of flood prediction as described below:

- A picture dataset is grouped into two subgroups. Among the two, one includes airborne images of flood-affected areas, while the other includes embossed pictures of regions which are not impacted by floods. The downloaded dataset is haphazardly apportioned from the outset and later the picture characterization is completed. This procedure comprises hybridization of SVM classifiers alongside k-mean clusters. Different classifications of the dataset are prepared utilizing the SVM classifier. We can see that SVM classifier is probably the best classifier amongst others that are generally utilized in remote detecting application and it gives a vastly improved precision on examination with different calculations like decision tree algorithms. Evaluation of the classifier depends on a few parameters, for example, precision, forecast speed and training time. Presentation of the classifier is presented utilizing the formed confusion matrix. This confusion matrix speaks about the quantity of accurately and inaccurately grouped models from the input data. Evaluation of this procedure is done after execution. It has the ability to classify the flooded regions with great precision. Almost all of the flooded pictures were viably portrayed by this prototype. The precision of the model is contrasted and examined via preparing it and various conditions and the outcomes are plotted.
- A constant checking system is prepared for evaluating parameters, for instance, precipitation power, soil dampness, level of water and pace of augmentation in the same. Various devices used for detection are fused into a structure, where information is logged and saved. The improvement of an estimate model reliant on fake neural framework, which have multiple layers, is made and is taken a stab at using a genuine arrangement. The response of the framework model which has multiple layers is approved by this evaluation. A little deviation regarding the real water level was seen in this prediction model, which is a significant issue as it prompts property harm, infrastructure damage and loss of human life too. In the progression of the estimate model, a multi-layered counterfeit neural framework model is used with the guide of MATLAB. In the genuine course of action, the framework is then modified and joined in the system. Model endorsement is done by running basics with unequivocal data sources and anticipated flood level as the output and it showed up to be diverse corresponding to the actual flood level after a foreordained range of time.

- Existing frameworks stick to tackle issues to forestall disastrous fiascos brought about by floods. The approach solidifies the flexibility of IOT and immovability of fake neural frameworks to manage the data given by a sensor device association and an early flood prediction is done by good correspondence between its two sections. It is exhibited that Levenberg Marquardt getting ready prediction with NARX orchestrate gives better lasting items and it similarly outfits constant flood prediction along with a phase ahead alert for genuine calamity to the officials. The occasion of a flash flood in an urban zone is anticipated a period phase-ahead and it is utilized to caution the nearby networks in danger of a looming fiasco. In this way, at times of viable crisis and calamity management, the board for smart countries is done by mistreating the most recent mechanical impedances, for example, IOT, ML, big data, farsighted assessment along with networked life and flexibility. The essential inspiration driving this structure is to screen the clamminess, pressure, temperature, precipitation, conduit level of water and to search their physical correspondent data for flood expectation and further examination. This IOT methodology is sent for data assortment from sensor-devices and completes its correspondence online and an ANN approach is used to investigate information recorded at the hour of flood forecast.

3. Proposed work

A flash flood is an immediate reaction to a place receiving precipitation having extremely high magnitude in a small timeframe. These sorts of floods are normally observed in urban regions of society where the primary land can't hold, or channel additional water elsewhere quickly by means of sewage frameworks and depleting waterways in a brief timeframe. As of late, the impact of floods in urban territories is highly visible in several cities, for example, Ahmedabad, Chennai, Mumbai and Trivandrum. Incorrect and delayed flood forecasting, poor urban planning, and inadequate flood moderation framework are the primary reasons behind this. Given dataset from various sources is used to form a generalized dataset, and applied to derive patterns and to get results with best efficiency. At this stage, the dataset will be used to stack in the information, check for tidiness, and afterwards, cut down and cleanse the given dataset for further examination. It is ensured that the given record steps cautiously and explain cleansing choices. The dataset used for analyzing given information is partitioned into a Training set and a Test set. Normally, 7:3 proportions are applied to perform this classification. The Data Framework which was formed using the Multi-Layer Perceptron Classifier will be applied to the Training set. Based on the test outcome and its accuracy, prediction of Test set is achieved.

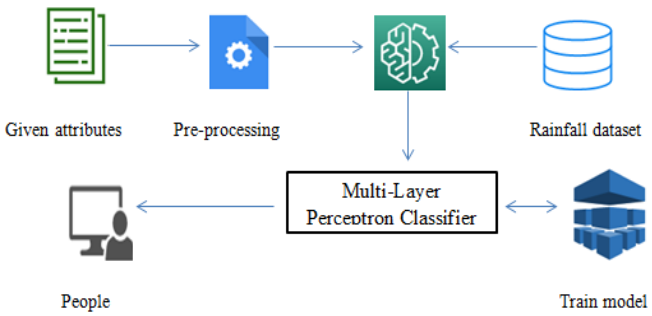

Figure 1. System Architecture

Among the various catastrophic events, floods are the most calamitous, reflecting enormous harm on human activity, foundations, agribusiness, and the financial framework. Political administrations, along these lines, feel the pressure to create solid, accurate and precise graphs identifying flood prone zones and additional arrangements for feasible and viable flood hazard controls concentrating on anticipation, preservation, and alertness. Flood forecasting systems are of high significance for risk evaluation and extraordinary event handling in case of an emergency. Robust, precise and correct forecasting contributes profoundly to water resource handling procedures, tactical strategies, further examination, and additional expulsion designing of the same. Hence, the significance of cutting-edge frameworks for short-period and long-period forecasting of flash floods and several other water-related blunders is emphatically asserted to reduce loss. Nonetheless, the forecasting of flood indication time and affected area is basically complex because of the changing nature of the atmospheric status.Existing frameworks stick to tackle issues to forestall disastrous blunders brought about by floods. The purpose of this system is to build a model for real-time flood prediction. In our proposed system, we took the best algorithms that would work and tested them to achieve good results. This system can predict flash floods in an urban area with better efficiency and accuracy than the conventional methods which are being used at present for the same. This proposed system helps us to implement better security and surveillance by predicting a flash flood well in advance and makes us ready to be able to tackle the same efficiently.

4. Implementation

Artificial Intelligence is sub-categorized into Machine Learning. The machine learning's objective is ideally to visualize the structure of data and consequently fit that data into models that can be inferred and used. Machine learning (ML) aids the computers in the ability to learn by itself rather than through programming. The different sorts of AI models incorporate supervised, unsupervised and reinforcement learning. The first type of learning algorithm, supervised learning, consists of both the input data and its corresponding learning data label, whose labelling is done manually. On the contrary, unsupervised learning has input data, but no labels. Supervised learning algorithms primarily deal with clustering of the input data. Reinforcement algorithms interact dynamically with its surrounding environment and then learn using responses which are either positive or negative.

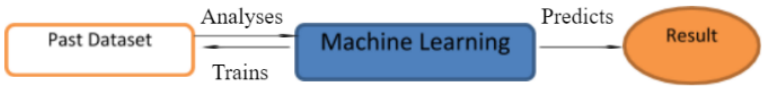

Fig: Process of Machine learning

Figure 2. Machine learning process

In supervised algorithms, one has multiple input factors (let=x_i, where i=1,2,3…..,n), a single output factor (let=y) and an algorithm for learning of the method that maps input, x, to the result, y, is f(x)=y. The purpose of the algorithm is to make the mapping function as analogous as possible so that when we have a new input data (x), the calculation will be able to predict the corresponding yield variable(y). The supervised algorithm consists of multiple methodologies for its implementation which are as follows: Support Vector Machines(SVMs), Multi-Class Classification, Logistic Regression, Decision Trees, etc. On the other hand, unsupervised learning algorithms which do not require any

administration as in the supervised model. Instead of that, the model itself is authorized to train on its own and thereafter identify the data. The unsupervised learning algorithms handle the unlabeled information. Although unsupervised learning has the ability to handle more compound problems, it is more unpredictable as compared to other algorithms.

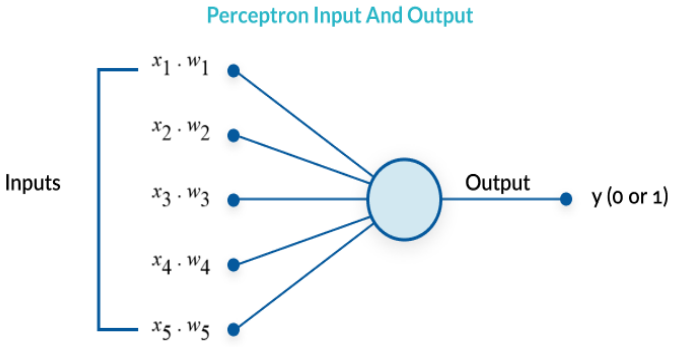

Figure 3. Perceptron input and output

Multi-layer Perceptron (MLP) belongs to the class of feed forward classifier of Artificial Neural Networks (ANN). Multi-Layer Perceptron(also called Multi-Layer Neural Network), contains one or more additional layers which are hidden layers(apart from single input and a single output layer). Whereas a single layer perceptron learns solely from the linear functions. On the other hand, a multi-layer perceptron has an added functionality to learn non–linear functions as well. There are at least 3 node layers which are as follows: output layer, input layer and hidden layer. Barring the input nodes, every other node is called a neuron which utilizes a nonlinear function for activation. The supervised learning procedure of backpropagation is utilized by MLP for the purpose of training our data-set.

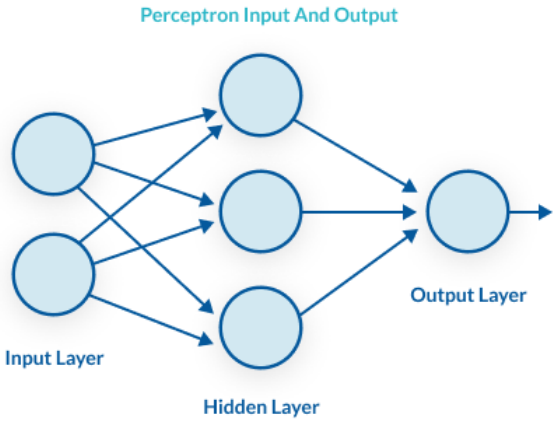

Figure 4. Multi-layer perceptron model

The modules included are as follows:

4.1.Data pre-processing

The task of pre-processing includes the alterations implemented to the data before dispensing it to the algorithm. It is a kind of preparation of data for the process. Raw data gets transformed into a clean data set using pre-processing. At whatever point the data is

aggregated from different sources, it is first accumulated in a raw format. This format isn't practical with the analysis process and the end goal of the probe. To accomplish better results from the applied model in the ML(Machine Learning) system for the data should be arranged in a proper manner. Some predetermined Machine Learning models need information in a predefined format; for example, Estimation of Random Forest algorithms does not contain values that are null or invalid. Along these lines, to implement random forest calculations, the values which are null or invalid must be taken care of or be removed from the first raw data-set.

4.2.Outlier detection process

Many AI calculations are delicate to the distribution and range of trait values in the dataset given as input. Exceptions in input information can alter and misguide the training procedure of AI calculations attaining higher training time, less accurate designs and finally impoverished outcomes.In truth, prior to anticipatory designs being arranged to assemble data, exceptions can generate deluding imitations and hence deceitful understandings of collected information. Anomalies can alter the outline dispersion of characteristic values in illustrative measurements like the average(mean) and SD(standard deviation) and in charts, for example, scatter plot graphs and histograms, compacting the body of data. Lastly, special cases can address cases of information instances that are significant to the issue, for example, peculiarities on account of detection of fraud, misrepresentation and computer security.

Outlier detection processes can neither befit the model on the training information nor assert that the design will function accurately for legitimate data. To achieve the aforementioned, we must ensure that the design has the correct models from the data, and is not finding a tremendous measure of clamor. Cross-approval is where we develop the model using the subset of the data-set and thereafter, evaluate using the correlative subset of the given data-set.

4.3.Data Visualization

Visualization of data is a significant ability in applied stats and AI. Applied stats do focus on large-scale depictions and assessment of data. Data recognition provides a noteworthy array of instruments for getting a subjective understanding of information. This can be useful when investigating and understanding a dataset and can help with recognizing designs or patterns in data, degenerate information, anomalies and substantially more. With a little domain intelligence, information perceptions may be utilized to direct and show key connections in plots and outlines that are more instinctive and partners than the proportions of association or significance. Information representation and exploratory information examination are entire fields themselves and it will suggest a more profound plunge into a portion of the books referenced towards the end.

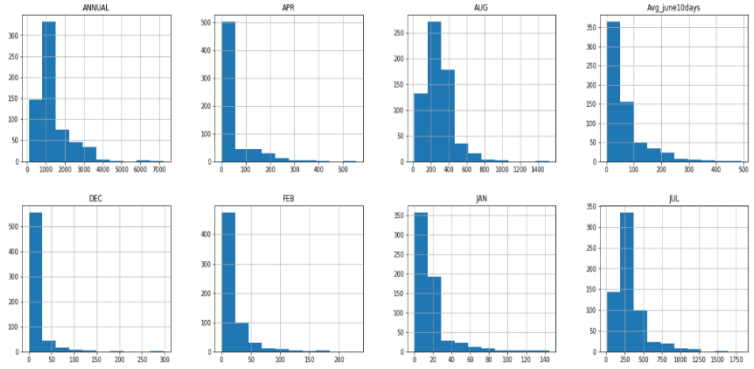

Figure 5. Rainfall distribution

In some cases, data does not bode well until it can be represented using visual forms or structures, for example, using outlines, charts and plots. Having the option to rapidly envision information samples is a significant ability both in applied stats and in applied machine learning. It will find the numerous kinds of plots that one should make note of while visualizing information in Python and how to utilize them to more readily understand one's own information.

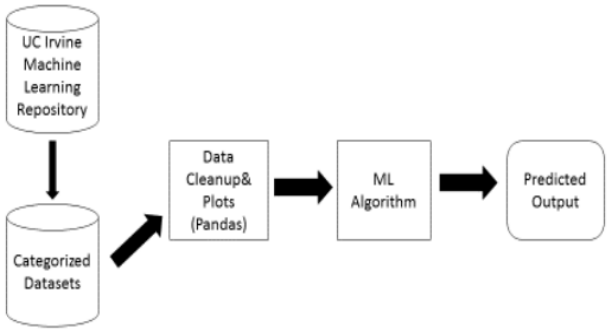

Figure 6. Stages of process

5. Results and Discussion

Upon running the program, the dataset can be plotted using the Matplotlib library and can be used to depict the monthly distribution of rainfall over a district of the Indian subcontinent. It helps to observe the months with maximum and minimum rainfall annually. The framework utilizes the normal precipitation of a third day of the month with moderate measures of precipitation received on a yearly premise. It can be improved by adding the effects of drainage and waterlogging systems and their efficiency in an urban area. The system can be improvised by adding a graphical user interface to make it user-friendly and a real-time application.

```
Classification report of MLP Classifier Results:

              precision    recall  f1-score   support

           0       0.97      1.00      0.99       185
           1       1.00      0.38      0.55         8

    accuracy                           0.97       193
   macro avg       0.99      0.69      0.77       193
weighted avg       0.97      0.97      0.97       193

Accuracy result of MLP Classifier is: 97.40932642487047

Confusion Matrix result of MLP Classifier is:
 [[185   0]
 [  5   3]]

Sensitivity :  1.0

Specificity :  0.375
```

Figure 7. MLP Classifier results

Machine Learning provides us with a platform to inhibit the system's ability to automatically learn and improve without being explicitly programmed. This process

ensures that any human error is omitted and a fool-proof mechanism is provided to get the desired results correctly over and over again.This project is based on the dataset of annual rainfall received in India over the last few years. It predicts the chances of occurrence of flash floods in an urban area of the society and helps prevent such disasters. The results provided by MLP classifier are the most accurate and provide better results in this field. The process of instant flood forecasting can be improvised in the time to come and help in enhancing the department for disaster management for the safety of humankind.

6. Conclusion

The proposed model will help us get rid of the obsolete model of manual entries and automation will help speed up the work. This model will help in recognizing and controlling flash floods in an urban area and also set risk management standards, based on the rainfall received in that region of the nation on a yearly basis. It lays focus on saving costs by being proactive instead of being reactive. It helps in the depreciation of the harm related to human survival, and decline in the property damage associated with floods.

Table 1.Comparison of results

S. No.	Algorithm used	Accuracy	Sensitivity	Specificity
1.	Logistic Regression	95.337	0.96	0.75
2.	Support Vector Classifier	95.855	1.0	0.0
3.	K-NearestNeighbour	95.855	0.98	0.375
4.	MLP Classifier	97.41	1.0	0.375

The future scope of this project is:

- Disaster management wants to automate the process of detection of flash floods in an urban area at real-time.
- To automate this process by displaying results of flood prediction in a web application or a desktop application.
- To optimize the work done by using and implementing an Artificial Intelligence environment.

References

[1] SurajOgale and Sanjay Srivastava, “Modelling and short-term forecasting of flash floods in an urban environment”, IEEE,**(2019)**.

[2] Akshya J and P. L. K. Priyadarsini, “A Hybrid Machine Learning Approach for Classifying Aerial Images of Flood-Hit Areas”, IEEE, **(2019)**.

[3] FebusReidj G. Cruz, Matthew G. Binag, Marlou Ryan G. Ga and Francis Aldrine A. Uy, “Flood Prediction Using Multi-Layer Artificial Neural Network in Monitoring System with Rain Gauge, Water Level, Soil Moisture Sensors”, IEEE, **(2018)**.

[4] Amitkumar, B. Ranit and Dr. P.V.Durge, “Different Techniques of Flood Forecasting and their Applications”, IEEE, **(2018)**.

[5] SwapnilBande and Prof. Dr. Virendra V. Shete, “Smart flood disaster prediction system using IoT & Neural Networks”, IEEE, **(2017)**.

[6] Gurleen Kaur and Anju Bala, "An Efficient Automated Hybrid Algorithm to Predict Floods in Cloud Environment", IEEE, **(2019)**.

[7] Nikhil Binoy C, Arjun N, Keerthi C, Sreerag S and Ashwin H Nair, "Flood Prediction using flow and depth measurement with Artificial Neural Network in canals", IEEE, **(2019)**.

[8] Ni KomangEga Kartika, Muhammad AryMurti and CasiSetianingsih, "Floods Prediction Using Radial Basis Function (RBF) Based on Internet of Things (IoT), IEEE, **(2019)**.

[9] Baskar.M, Gnansekaran.T(2017), "Multi Model Network Analysis for Improved Intrusion Tracing towards Mitigating DDoS Attack", Asian Journal of Research in Social Sciences and Humanities, Vol.7, No.3, pp.1343-1353.

[10] Baskar.M, Gnansekaran.T(2017),"Developing Efficient Intrusion Tracking System using Region Based Traffic Impact Measure Towards the Denial of Service Attack Mitigation", Journal of Computational and Theoretical Nanoscience, Vol.14, No.7, pp: 3576-3582.