

FLOOD PREDICTION USING MACHINE LEARNING

¹Dr. L. Nalini Joseph, ²K. Sai Sateesh, ³K. Siva Prasad, ⁴T. Ram Priyatham

^{1,2,3,4}Professor

Dept of Computer Science and Engineering,
Bharath Institute of Higher Education and Research
Chennai, Tamilnadu, India

Abstract- Weather forecasting has historically been performed with air frame fashions, which are inconsistent with disturbances and consequently had been erroneous over long periods of time. Because machine mastering methods are more prone to perturbations, in this article we explore their application to weather forecasting to doubtlessly generate extra correct weather forecasts over longer periods. The scope of this article is limited to predicting the maximum and minimal temperatures for seven days, contemplating the weather records of the closing two days. A linear regression model and a variance regression version are used to capture weather tendencies. The essential goal of this dissertation turned into to predict rainfall using mining techniques. The purpose of this dissertation changed into to tell farmers approximately the possibility of rains a good deal in advance. As a end result, there may be fewer troubles. Basically, this thesis will make contributions to the agricultural sector. Every yr a few areas are submerged below water because of the rains. As a result, agricultural fields are badly broken and lots of herds of cattle suffer from this. Many homes have additionally fallen into disrepair due to storms and floods that affect each day life. Most of the victims are human beings dwelling at the banks of the river. Not simplest the population of these areas suffer, however additionally different villages, as such situations result in a meals crisis. This end becomes finished using device studying algorithms.

INTRODUCTION

Predicting a flood is a totally hard task. Although many algorithms have already been proposed, it is nonetheless very difficult to predict rainfall appropriately. In an agrarian u .S . Like India, the achievement or failure of the harvest and the shortage of water in each year is constantly regarded upon with amazing challenge. Small fluctuations inside the seasonal rains can wreak havoc in an agricultural vicinity. Accurate precipitation forecasting has the ability to help save you the causation and damage as a result of natural disasters. In some instances, including floods and droughts, a totally accurate forecast of rainfall is useful for agricultural management and catastrophe prevention. Various algorithms are explained in this newsletter. Data mining strategies are effectively utilized in precipitation forecasting.

OBJECTIVES

The motive of flood forecasting using the document is to broaden a version for flood forecasting the usage of rainfall records. Predictions from one of a kind fashions are taken and compared to every other to locate the first-rate model with the best accuracy. Flood in different states of India can be anticipated in distinct months. The confusion matrix of diverse models in system mastering is considered to evaluate the accuracy and precision of the system.

EXISTING SYSTEM

Our agriculture is the energy of the Indian financial system. The farmer's cultivation time relies upon simplest at the rainy season. Good vegetation handiest require top fertilizer and properly climate. A weather forecast is the maximum important requirement of every farmer. Due to unexpected climate/climate modifications, human beings are struggling economically and physically. Predictable weather is one of the maximum complicated troubles inside the cutting-edge world. The important reason of this text is to expect the weather the use of diverse facts mining strategies. Such as class, clustering, choice tree, and neural networks. Weather-related data is also referred to as meteorological facts. In this newsletter, the maximum common weather parameters used are precipitation, wind velocity, temperature, and cold.

DISADVANTAGES OF THE EXISTING SYSTEM

- Classification
- Clustering
- Decision tree

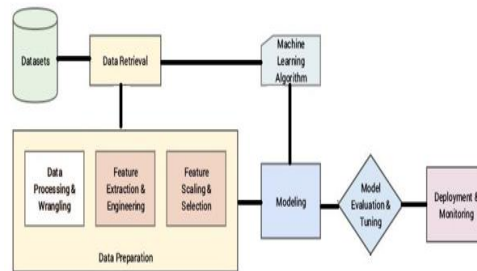
PROPOSED SYSTEM

- Flooding is vital for the food production plan, water management and all nature motion plans.
- The presence of extended dry spells or heavy rainfall all through vital durations of plant boom and improvement can cause significant yield reductions.
- India is an agricultural united states and its financial system is basically primarily based on crop yields.
- Thus, precipitation forecasting becomes very important in agricultural regions like India.
- Flood forecasting has been one of the finest clinical and technical demanding situations of the past century international.

ADVANTAGES OF PROPOSED SYSTEM

- Weather forecast numbers
- Statistical climate forecast
- Synoptic weather forecast

SYSTEM ARCHITECTURE



SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS:

- System - Windows 7/10
- Speed - 2.4GHZ
- Hard disk - 40GB
- Monitor - 15VGA color
- RAM - 4GB

SOFTWARE REQUIREMENTS:

- Operating System - Windows XP
- Coding language - PYTHON
- IDE - Jupyter Notebook(Anaconda)

LITERATURE ANALYSIS

A. Indian Summer Monsoon Rainfall (ISMR) Forecasting using Time Series Data: A Fuzzy-Entropy-Neuro based Expert System In [30],

The authors presented a version for predicting summer time monsoon rainfall at month-to-month and seasonal scales. The forecast usage time collection facts covers the length from 1871 to 2014. The dataset become break up into two elements, (1) 1871-1960 turned into used as schooling facts and (2) 1961-2014 changed into used as test statistics. Statistical evaluation found out a dynamic rainfall sample in the course of the wet season that might be efficiently predicted with the aid of mathematical and statistical models. Therefore, the authors in this observe endorse the usage of three techniques for this type of prediction: fuzzy set, entropy, and artificial neural network. Using these 3 techniques, a predictive version is advanced that considers the dynamic nature of ISMR. In the proposed version, although, set idea is used to handle the uncertainties which can be inherent in certain records. The idea of entropy calculation become changed in this model and used to provide enter for the degree of club within the entropy function. This characteristic is known as entropy although the information advantage (FIG). Then each of them using the smoky management of ANN. The FIG price of every parameter set became then used as enter to the ANN. The proposed model is named "Fuzzy Entropy and Neural Network Expert System for ISMR Prediction" because it's far an integration of any goal, entropy and ANN. To evaluate the proposed version performance, the subsequent measures of accuracy have been used: trendy errors (SD), correlation coefficient (CC), root imply square errors (RMSE), and performance parameters (PP). According to the results, the proposed version is efficient and within your means as compared to different existing fashions.

B. An Extensive Evaluation of Seven Machine Learning Methods for Rainfall Prediction in Weather Derivatives

The researchers compared the predictive overall performance of the modern day and modern day technique called "Markov Chain Augmented with Precipitation Prediction" with other broadly used getting to know tools: vector regression, genetic programming, M5 rules, M5 model timber, radials. Neural network Networks and okay acquaintances. Daily rainfall video data were gathered from 42 towns throughout two continents with special weather structures. 20 European towns and 22 US towns had been selected. There are motives for deciding on continents to extract records: one to check one of a kind climates and unique weathers, the opposite due to localities, whilst the chosen towns are a ways apart from each other. The last aim turned into now not to tie the experiment to a particular type of weather or to a selected geographic vicinity. Depending on the outcomes, the accumulation of precipitation can provide top results for forecasting the usage of daily rainfall statistics. Support vector regression, radial basis features, and genetic programming typically carried out properly with backlog statistics, however radial foundation functions done better than contemporary Markov chain strategies. For all decided on records, each method used the same parameters, so there has been no guarantee that the excellent parameter setting became used for all methods. Through the test, the researchers referred to the

relationship between the accuracy and characteristics of climatic forecasting, along with: the variable nature of precipitation, the maximum amount of precipitation, and the interquartile distance of precipitation. At the equal time, there was no huge difference within the prediction blunders of the algorithms among the states of each continents (USA and Europe). The hassle of inhomogeneity in precipitation data become solved with the assist of cumulative precipitation sums.

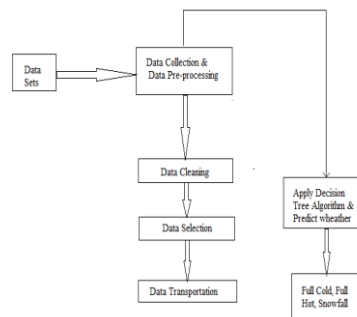
C. A Hybrid Model for Statistical Downscaling of Daily Rainfall

The authors proposed a hybrid method of lowering every day precipitation by integrating two techniques: 1) random forests and 2) protection vector machines. RF become selected because of its robustness in type and was used to are expecting whether or not it would rain or not, at the same time as SVM changed into selected because of its property of fitting non-linear information and become used to predict the amount of rain if it did occur. . The proposed version become evaluated via downscaling daily precipitation at three stations: Dungun, Besut and Kemaman at the jap coast of Peninsular Malaysia. Time collection facts for every day precipitation from 1961 to 2000 turned into acquired from the Department of Irrigation and Drainage Malaysia. A general of 26 weather facts have been accrued through the National Center for Environmental Prediction reanalysis datasets, which were used for upscaling models. Various fine control operations had been performed to evaluate the homogeneity of the precipitation time series. Histograms of the dataset had been created to mirror the questions, further, the pupil's t-take a look at was also used to discover differences within the way among the 2 parts of the dataset, which ended up being uniform in all 3 areas. According to the outcomes, the hybrid method can reduce precipitation with a Nash-Sutcliff performance in the variety of zero.90-0.93, that's appreciably better than the RF and SVM models.

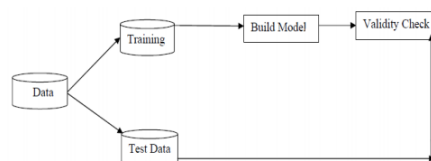
DATA FLOW DIAGRAM

1. A DFD is also referred to as a bubble chart. It is a simple graphical formalism that can be used to represent a device in terms of inputs to the gadget, the various tactics accomplished on that information, and the outputs generated with the aid of it.
2. Data float diagram (DFD) is one of the principal modeling equipment. It is used to model elements of the system. These components are the system strategies, the information utilized by the technique, the outside item that corresponds to the machine, and the records flows inside the machine.
3. The DFD shows how records movements via the gadget and how it's far changed through a sequence of changes. It is a graphical method that depicts the float of facts and the modifications that are carried out to transport the statistics from enter to output.
4. A DFD is likewise known as a bubble chart. A DFD can be used to represent a system at any degree of abstraction. A DFD can be divided into layers that constitute incremental facts glide and person operations.

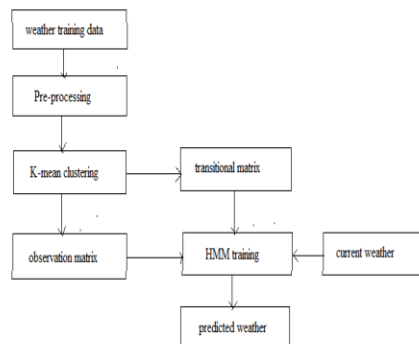
LEVEL - 0



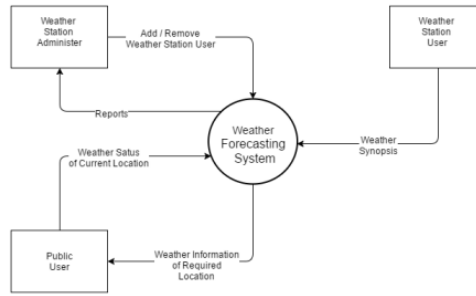
LEVEL - 1



LEVEL - 2



LEVEL – 3



UML DIAGRAMS

UML stands for Code of Canon Law. UML is a wellknown cause modeling language for item-oriented software program development. The flag is controlled and created by the item control group.

UML is intended to end up a common language for developing item-oriented computer software models. In its modern shape, UML has two primary components: the metamodel and the notation. Certain methods or sorts of approaches will also be delivered in the future; or to the UML.

The Unified Modeling Language is a wellknown language for expressing, visualizing, building, and documenting the structure of software program structures, as well as for modeling business and other non-software systems.

UML Sets engineering fine practices which have demonstrated to be powerful in modeling massive and complicated systems.

UML is an crucial a part of object-orientated software program improvement and the software development process. UML specifically makes use of graphical notation to layout software program projects.

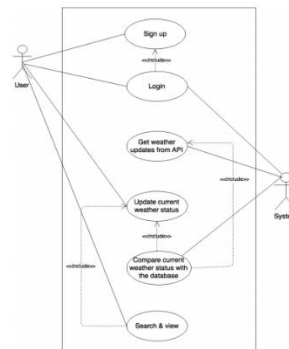
GOALS:

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extensibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modeling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

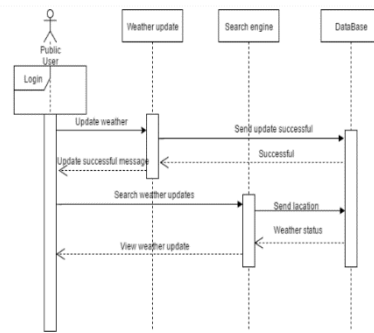
USE CASE DIAGRAM:

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.



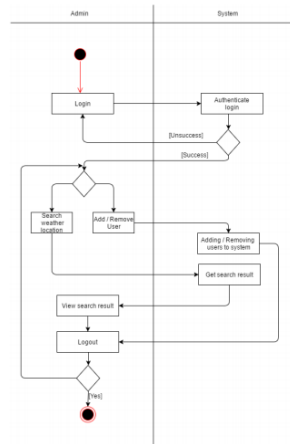
SEQUENCE DIAGRAM:

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams



ACTIVITY DIAGRAM:

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of contro



REFERENCES:

1. Xiong, Lihua, and Kieran M. OConnor. "An empirical method to improve the prediction limits of the GLUE methodology in rainfallrunoff modeling." *Journal of Hydrology* 349.1-2 (2008): 115-124.
2. Schmitz, G. H., and J. Cullmann. "PAI-OFF: A new proposal for online flood forecasting in flash flood prone catchments." *Journal of hydrology* 360.1-4 (2008): 1-14.
3. Riordan, Denis, and Bjarne K. Hansen. "A fuzzy casebased system for weather prediction." *Engineering Intelligent Systems for Electrical Engineering and Communications* 10.3 (2002): 139-146.
4. Guhathakurta, P. "Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network." *Current Science* 90.6 (2006): 773-779.
5. Pilgrim, D. H., T. G. Chapman, and D. G. Doran. "Problems of rainfall-runoff modelling in arid and semiarid regions." *Hydrological Sciences Journal* 33.4 (1988): 379-400.
6. Lee, Sunyoung, Sungzoon Cho, and Patrick M. Wong. "Rainfall prediction using artificial neural networks." *journal of geographic information and Decision Analysis* 2.2 (1998): 233-242..
7. French, Mark N., Witold F. Krajewski, and Robert R. Cuykendall. "Rainfall forecasting in space and time using a neural network." *Journal of hydrology* 137.1-4 (1992): 1-31.
8. Charaniya, Nizar Ali, and Sanjay V. Dudul. "Committee of artificial neural networks for monthly rainfall prediction using wavelet transform." *Business, Engineering and Industrial Applications (ICBEIA), 2011 International Conference on. IEEE, 2011.*
9. Noone, David, and Harvey Stern. "Verification of rainfall forecasts from the Australian Bureau of Meteorology's Global Assimilation and Prognosis(GASP) system." *Australian Meteorological Magazine* 44.4 (1995): 275-286.
10. Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators." *Neural networks* 2.5 (1989): 359-366.
11. Haykin, Simon. *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.
12. Rajeevan, M., PulakGuhathakurta, and V. Thapliyal. "New models for long range forecasts of summer monsoon rainfall over North West and Peninsular India." *Meteorology and Atmospheric Physics* 73.3-4 (2000): 211-225.