

TABLE OF CONTENTS

Chapter	Title Name	pg.no
	Abstract	3
1	Introduction	4
2	Literature Survey	5
3	Aim and scope of present investigation	7
	3.1:Module Implementation	7
	3.2:Data flow Diagrams	8
	3.3:UML Diagrams	10
4	Experimental or Scope of the present investigation Algorithm used	13
5	Results and discussion, performance analysis	15
6	Summary and Conclusion	19
	Plagiarism Report	20

RAINFALL PREDICTION WITH AGRICULTURAL SOIL ANALYSIS USING MACHINE LEARNING

Abstract

India is an agricultural country and its economy is largely based upon crop productivity and rainfall. For analyzing the crop productivity, rainfall prediction is require and necessary to all farmers. Rainfall Prediction is the application of science and technology to predict the state of the atmosphere. It is important to exactly determine the rainfall for effective use of water resources, crop productivity and pre planning of water structures. Using different data mining techniques it can predict rainfall. Data mining techniques are used to estimate the rainfall numerically. This paper focuses some of the popular data mining algorithms for rainfall prediction. Naive Bayes, K-Nearest Neighbour algorithm, Decision Tree are some of the algorithms compared in this paper. From that comparison, it can analyze which method gives better accuracy for rainfall prediction. Prediction of crops may be accurately through with the help of data mining techniques and considering the environmental parameters. During this work, the classifiers used area unit support vector machine and data processing.

CHAPTER 1

INTRODUCTION

Rainfall Prediction is one of the most challenging tasks. Though already many algorithms have being proposed but still accurate prediction of rainfall is very difficult. In an agricultural country like India, the success or failure of the crops and water scarcity in any year is always viewed with greatest concern. A small fluctuation in the seasonal rainfall can have devastating impacts on agriculture sector. Accurate rainfall prediction has a potential benefit of preventing causalities and damages caused by natural disasters. Under certain circumstances such as flood and drought, highly accurate rainfall prediction is useful for agriculture management and disaster prevention. In this paper, various algorithms have been analyzed. Data mining techniques are efficiently used in rainfall prediction

CHAPTER 2

LITERATURE SURVEY

Indian Summer Monsoon Rainfall (ISMR) Forecasting using Time Series Data: A Fuzzy-Entropy-Neuro based Expert System Pritpal Singh / 30 July 2018 Statistical analyzes reflect the dynamic nature of the ISMR, which couldn't be predicted efficiently by statistical and mathematical based models. Therefore, this study suggests the usage of three techniques, viz., fuzzy set, entropy and artificial neural network (ANN). Based on these techniques, a novel ISMR time series forecasting model is designed to deal with the dynamic nature of the ISMR. This model is verified and validated with training and testing data sets. Various statistical analyzes and comparison studies demonstrate the Less Severity.

An Extensive Evaluation of Seven Machine Learning Methods for Rainfall Prediction in Weather Derivatives Sam Cramer a,* , Michael Kampouridis a , Alex A. Freitas a , Antonis K. Alexandridis b / 2017 This work's main impact is to show the benefit machine learning algorithms, and more broadly intelligent systems have over the current state-of-the-art techniques for rainfall prediction within rainfall derivatives. We apply and compare the predictive performance of the current state-of-the-art (Markov chain extended with rainfall prediction) and six other popular machine learning algorithms one of the disadvantages of MCRP is that MCRP is purely driven off the historical data of each city and needs to be tuned according to the specific

A Hybrid Model for Statistical Downscaling of Daily Rainfall Sahar Hadi Poura , Shamsuddin Shahida , Eun-Sung Chungb / The RF was used to predict whether rain will occur in a day or not and SVM was used to predict amount of rainfall in rainfall occurring days. The capability of proposed hybrid model was verified by downscaling daily rainfall at three rain-gauge locations in the east coast of peninsular Malaysia. The hybrid model was also found to replicate the variability, number of consecutive wet days, 95-percentile rainfall amount in each months as well as distribution of SVM considers a kernel-based ANN to overcome the disadvantages of the traditional ANN and therefore, able to develop relationship among highly non-linear variables

: A REVIEW 2021 In India ,Agriculture is the key point for survival. For agriculture, rainfall is most important. These days rainfall prediction has become a major problem. Prediction of rainfall gives awareness to people and know in advance about rainfall to take certain precautions to protect their crop from rainfall. Many techniques came into existence to predict rainfall. Machine Learning algorithms are mostly useful in Forecasts are never 100% accurate.

CHAPTER 3

Aim and Scope of the present investigation

3.1: Module Implementation

- Data Collection
- Data Cleaning
- Data Selection
- Data Transformation
- Data Mining Stage

Data Collection

The data used for this work was collected from meteorologist's centre. The case data covered the period of 2012 to 2015. The following procedures were adopted at this stage of the research: Data Cleaning, Data Selection, Data Transformation and Data Mining.

Data Cleaning

In this stage, a consistent format for the data model was developed which is search missing data, finding duplicated data, and weeding out of bad data. Finally system cleaned data were transformed into a format suitable for data mining

Data Selection

At this stage, data relevant to the analysis like decision tree was decided on and retrieved from the dataset. The Meteorological dataset had ten attributes in that were using two attributes for future prediction. Due to the nature of the Cloud Form data where all the values are the same and the high percentage of missing values in the sunshine data both were not used in the analysis.

Data Transformation

“This is also known as data consolidation”. It is the stage in which the selected data is transformed into forms appropriate for data mining. The data file was saved in Commas Separated Value (CVS) file format and the datasets were normalized to reduce the effect of scaling on the data.

Data Mining Stage

The data mining stage was divided into three phases. At each phase all the algorithms were used to analyse the meteorological datasets. The testing method adopted for this research was percentage split that train on a percentage of the dataset, cross validate on it and test on the remaining percentage. There after interesting patterns representing knowledge were identified

3.2: DATA FLOW DIAGRAM

LEVEL - 0

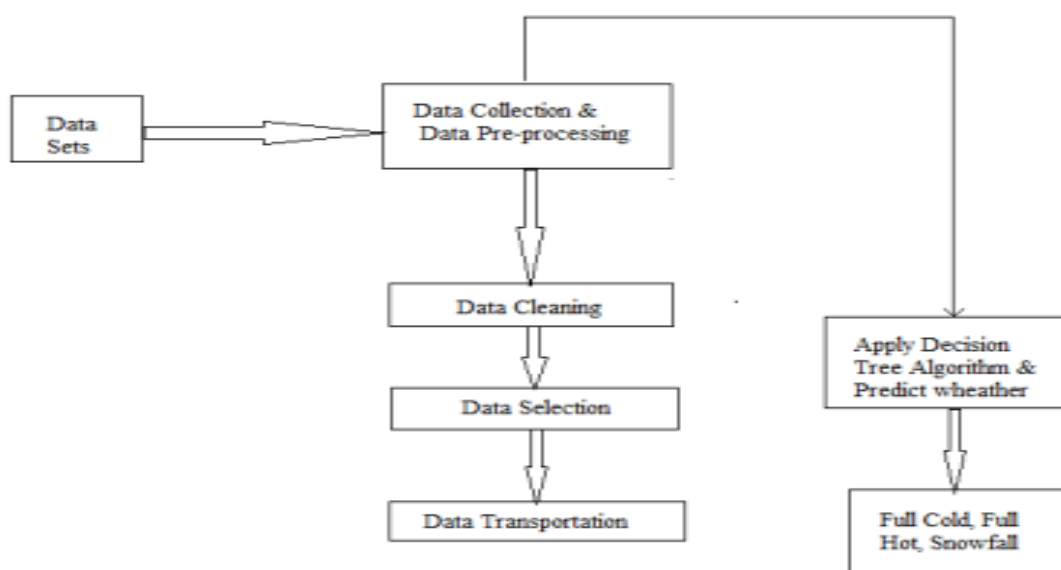


Fig a: Data Flow Diagram

LEVEL - 1

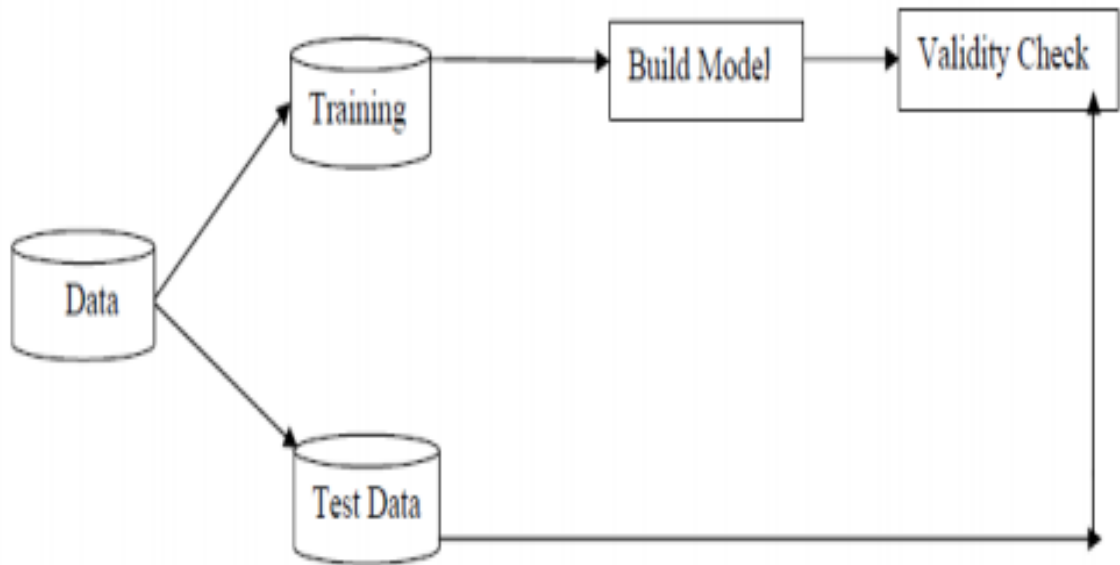


Fig a.1: Level 1

LEVEL2

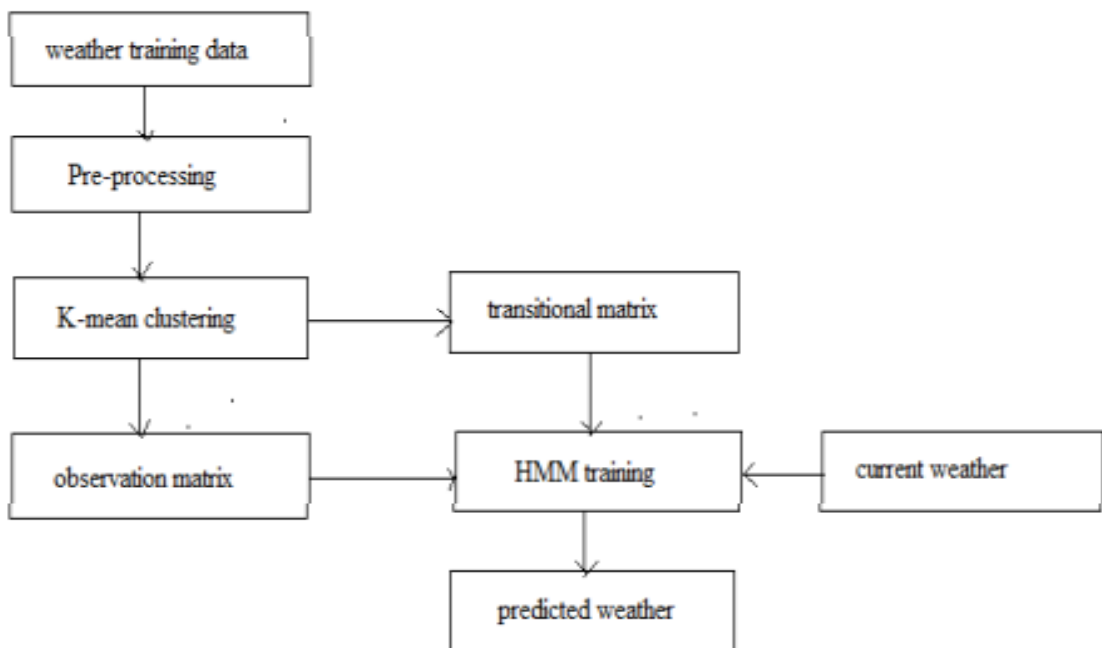


Fig a.2: Level 2

3.3: UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group. The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML. The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems. The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems. The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

GOALS:

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

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

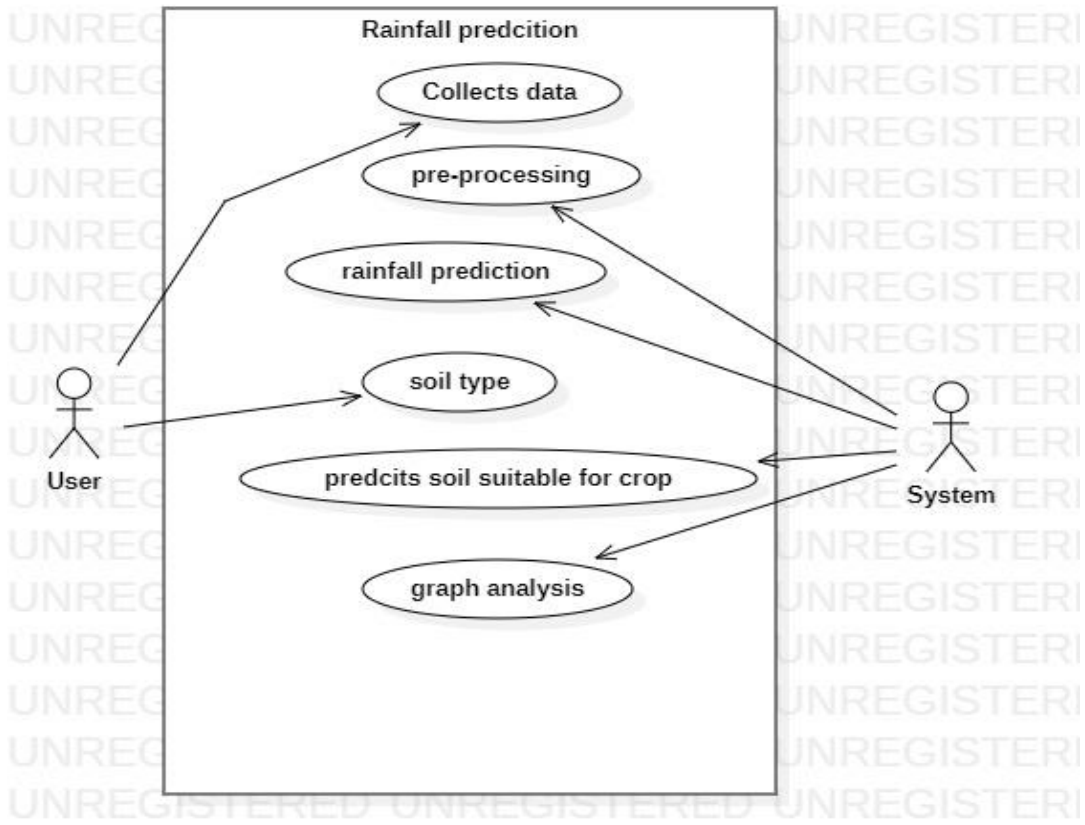


Fig b: Usecase Diagram

CHAPTER 4

Experimental or Scope of the present investigation Algorithm used

- Rainfall is important for food production plan, water resource management and all activity plans in the nature.
- The occurrence of prolonged dry period or heavy rain at the critical stages of the crop growth and development may lead to significant reduce crop yield.
- India is an agricultural country and its economy is largely based upon crop productivity.
- Thus rainfall prediction becomes a significant factor in agricultural countries like India.

Rainfall forecasting has been one of the most scientifically and technologically challenging problems around the world in the last century.

System Architecture

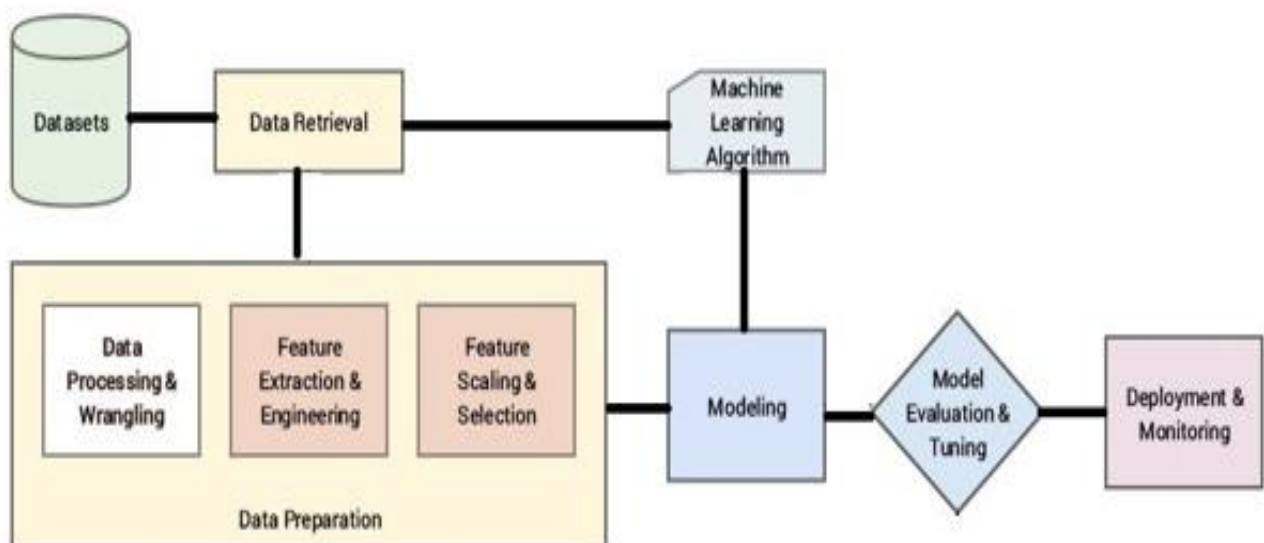


Fig d: System Architecture