

A NOVEL APPROACH FOR PREDICTION OF DROUGHT

Submitted in partial fulfillment of the requirements for
the award of
Bachelor of Engineering degree in Computer Science and Engineering

by

ELURI HARICHANDANA (Reg. No. 37110208)

GUDIPUDI RAMYA(Reg. No.37110248)



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SCHOOL OF COMPUTING**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

**Accredited with Grade "A" by NAAC
JEPPIAAR NAGAR, RAJIV GANDHI SALAI,
CHENNAI – 600 119**

MARCH - 2021



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

Accredited with "A" grade by NAAC

Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600 119

www.sathyabama.ac.in



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this project report is the bonafide work **ELURI HARICHANDANA (Reg. No. 37110208) and GUDIPUDI.RAMYA (Reg No. 37110248)** who carried out the project entitled "**A NOVEL APPROACH FOR PREDICTION OF DROUGHT**" under my supervision from **November 2020** to **March 2021**.

Internal Guide

Dr.T.Sasikala, M.E.,Ph.D.,

Head of the Department

Submitted for Viva voce Examination held on _____

Internal Examiner

External Examiner

DECLARATION

I **ELURI HARICHANDANA** and **GUDIPUDI RAMYA** hereby declare that the Project Report entitled “**A NOVEL APPROACH FOR PREDICTION OF DROUGHT**” is done by me and my partner under the guidance of **Dr.T.Sasikala, M.E.,Ph.D.**,Department of Computer Science and Engineering at Sathyabama Institute of Science and Technology is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Computer Science and Engineering.

DATE:

PLACE:CHENNAI

SIGNATURE OF THECANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management** of **SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T. Sasikala, M.E., Ph.D., Dean**, School of Computing, **Dr. S. Vigneswari, M.E., Ph.D.,** and **Dr. L. Lakshmanan, M.E., Ph.D., Heads of the Department** of Computer Science and Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide, **Dr. T. Sasikala, M.E., Ph.D., Dean** for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Computer Science and Engineering** who were helpful in many ways for the completion of the project.

ABSTRACT

One of the major challenges of agricultural systems is how to mitigate the impacts of droughts. To mitigate the impacts of drought, an effective and timely monitoring system is required. Effective monitoring of droughts can aid in developing an early warning system. Data mining extracts the valuable information from such large data archives in an automated and efficient way. Most of the predictions primarily depend on precipitation rate. Precipitation rate is a major factor in causing drought, but when predicting for a localized area, there are certain other factors which can also have a potential impact on drought. So, this methodology involves basic parameters like humidity, temperature, ground water level, rainfall, wind speed for predicting drought. This is also a less complex method in predicting drought, comparing with existing drought predicting methodologies.

TABLE OF CONTENTS

	ABSTRACT	v
	LIST OF FIGURES	viii
CHAPTER No.	TITLE	PAGE No.
1.	INTRODUCTION	1
	1.1 PROJECT OVERVIEW	1
	1.2 METHODS FOR PREDICTING DROUGHT OCCURANCES	2
	1.3 AGRICULTURAL DROUGHT FORECASTING AND WARNING SYSTEM	5
2.	AIM AND SCOPE	7
	2.1 AIM OF PROJECT	7
	2.1.1 <i>OBJECTIVES</i>	7
	2.2 SCOPES	7
3.	SYSTEM DESIGN & METHODOLOGY	12
	3.1 PROPOSED SYSTEM	12
	3.2 EXISTING SYSTEM	12
	3.3 SYSTEM CAPABILITIES	13
	3.4 REQUIREMENT SPECIFICATION	13
	3.5 PREDICTING OF DROUGHT AREAS	14
	3.6 DROUGHT PREDICT IN DIFFERENT AREAS	18
	3.7 MODULES	23
	3.7.1 <i>ANALYSIS MODULE</i>	23
	3.7.2 <i>DROUGHT PREDICTION MODULE</i>	23

3.7.3	<i>CALCULATIVE SEVERITY INDEX</i>	24
3.7.4	<i>VIEW INFORMATION MODULE</i>	24
4.	SOFTWARE DEVELOPMENT METHODOLOGY	25
4.1	DESCRIPTION OF DIAGRAM	25
4.2	ARCHITECTURAL DESIGN	26
4.3	SEQUENCE DIAGRAM	27
4.4	USE CASE	27
4.5	COLLABORATION DIAGRAM	28
4.6	ACTIVITY DIAGRAM	29
5.	RESULTS AND DISCUSSION	30
6.	CONCLUSION AND FUTURE WORK	35
	REFERENCES	37
	APPENDIX	
A.	PAPER ACCEPTANCE MAIL	38
B.	JOURNAL PAPER	39
C.	SIGNED BONAFIDE	42
D.	SOURCE CODE	43

LIST OF FIGURES

FIGURE No.	FIGURE NAME	PAGE No.
4.1	BLOCK DIAGRAM	25
4.2	ARCHITECTURAL DIAGRAM	26
4.3	SEQUENCE CASE	27
4.4	USE CASE DIAGRAM	27
4.5	COLLABORATION DIAGRAM	28
4.6	ACTIVITY DIAGRAM	29
5.1	TO RUN THE PROGRAM	30
5.2	ANACONDA COMMAND PROMPT	31
5.3	RUNNING AS JAVA APPLICATION	32
5.4	DROUGHT DOES NOT OCCUR	33
5.5	DROUGHT OCCUR	34

CHAPTER 1

INTRODUCTION

1.1 PROJECT OVERVIEW

A drought is a prolonged shortages in the water supply, whether atmospheric, ground water or surface water in an area. It can substantially affect the biological community of the influenced area and damage to the nearby ecosystem. Continuous drought and water levels below the minimum leading to direct danger for amphibian life. Drought is an adverse environmental phenomenon that influences almost all aspects of society.

It is one of the features of climate and its occurrence is unavoidable. Among all extreme climate events, drought is the most complex phenomenon affecting more people. Identifying past droughts and predicting future ones is very vital in limiting their effects. However, the random and nonlinear nature of drought variables makes accurate drought prediction remain a challenging scientific problem. This method predicts drought by assessing impact of individual factors on drought, analyzing the relationship in them and draws the conclusion regarding the drought.

Medium range soil moisture drought forecast helps in decision making in the field of agriculture and water resources management. Part of skills in medium range drought forecast comes from precipitation. Proper evaluation and correction of precipitation forecast may improve drought predictions. Here, we evaluate skills of ensemble mean precipitation forecast from Global Ensemble Forecast System (GEFS) for medium range drought predictions over India. Climatological mean (CLIM) of historic data (OBS) are used as reference forecast to evaluate GEFS precipitation forecast.

Analysis was conducted based on forecast initiated on 1st and 15th dates of each month for lead up to 7-days. Correlation and RMSE were used to estimate skill scores of accumulated GEFS precipitation forecast from lead 1 to 7-days. Volumetric indices based on the 2X2 contingency table were used to check missed and falsely predicted historic volume of daily precipitation from GEFS in different regions and at different thresholds. GEFS showed improvement in correlation of 0.44 over CLIM during the monsoon season and 0.55 during the winter season.

Lower RMSE was showed by GEFS than CLIM. Ratio of RMSE in GEFS and CLIM comes out as 0.82 and 0.4 (perfect skill is at zero) during the monsoon and winter season, respectively. We finally used corrected GEFS forecast to derive the Variable Infiltration Capacity (VIC) model, which was used to develop short-term forecast of hydrologic and agricultural (soil moisture) droughts in India.

The current NMME SPI prediction framework is similar to the CFSv2 SPI prediction system that developed by Yoon et al. (2012). For each model, monthly-mean precipitation (P) forecasts were first bias corrected and spatially downscaled (BCSD) to regional grids of 0.5-degree resolution over the contiguous United States based on the probability distribution functions (PDFs) derived from the hindcasts.

As a result, BCSD scheme corrects both the climatological mean and standard deviation of the hindcasts. Specifically, for each month and lead time, the PDF at each grid point is computed based on model hindcasts excluding the target year. The bias-corrected percentile for the target year is then obtained from the inverse PDF of the P analysis based on the percentile calculated from the PDF of the hindcasts. The BCSD method was applied to each member and each lead of the P hindcasts. The corrected P forecasts were then appended to the CPC Unified Precipitation Analysis (Chen et al. 2008) to form a P time series for computing 1-month, 3-month, 6- month, and 12-month SPIs.

The NMME-ensemble SPI forecasts are the equally weighted mean of the six model forecasts. Two performance measures, the anomaly correlation coefficient (ACC) and root-meansquare errors (RMSE) against the observations, are used to evaluate forecast skill. In this study, CPC Unified Precipitation Analysis is used as the observations for forecast evaluation. Analyzes climate of each location and adjusts duration coefficients (empirical factors) accordingly Duration factors affect the sensitivity of the index to moisture deficits by determining how much weight is given to the current moisture anomaly and the previous index value Two sets of duration coefficients are calculated for every location (one for wet spells, other for dry spells).

climate characteristic coefficient calibration: 2 % of PDSI values fall below - 4.0 and 2 % rise above 4.0 => this gives index an uper and lower bounds 2 % of extreme dry events and 2 % of extreme wet events Self calibrated PDSI performs much more consistently than the original PDSI (Wells et al., 2004)* and provides more realistic metric of relative periods of drought or excessive soil moisture.

1.2 METHODS FOR PREDICTING DROUGHT OCCURANCES

A comprehensive assessment of future dry events in a region is essential for finding sustainable solutions for water-related problems concerning water management and preventive risk assessment of drought. It is also a first step in developing a common European strategy for facing up to the negative impacts of present and future dry climatic events. A useful index for drought monitoring, based only on monthly precipitation, is the Standardized Precipitation Index (SPI); here we assume that it properly describes the climatic condition of a particular region.

By applying an appropriate forecast method to the precipitation time series and then computing the SPI, it is possible to predict future drought occurrences. Forecasting time series is a well-known field of Statistics and methods are well documented in the scientific literature. In the present paper we apply a standard technique and we contrast it with a new method here presented. As a test case, we use rain gauge data for Sicily, which may be considered a key region for understanding climatic conditions in the Mediterranean basin. Drought forecasting and monitoring can be defined as data collection, analysis, and reporting to develop information on current and future drought conditions.

Observing and understanding current and historical weather, soil, and hydrologic conditions during droughts goes hand-in-hand with predicting future droughts, as these observations are the main inputs for computer models used in drought forecasting. Drought reports derived from monitoring data and forecasts simulated in computer models typically report drought in terms of the extent, severity, and duration of drought conditions.

There is not a universally accepted way of measuring/forecasting drought; however the various measures used to create maps are largely calculated from temperature, precipitation, streamflow, and soil moisture data. Canada's official drought reports (Canadian Drought Monitor, CDM) rely on multiple calculations and datasets, including satellite imagery and well-established indices, such as the Palmer Drought Severity Index (PDSI) and the Standardized Precipitation Index (SPI).

The CDM is a map of current drought conditions while forecasts use the same calculations and include future inputs. This paper summarizes and synthesizes the research carried out under the NOAA Drought Task Force (DTF) and submitted in this special collection. The DTF is organized and supported by NOAA's Climate Program Office with the National Integrated Drought Information System (NIDIS) and involves scientists from across NOAA, academia, and other agencies.

The synthesis includes an assessment of successes and remaining challenges in monitoring and prediction capabilities, as well as a perspective of the current understanding of North American drought and key research gaps.

Results from the DTF papers indicate that key successes for drought monitoring include the application of modern land surface hydrological models that can be used for objective drought analysis, including extended retrospective forcing datasets to support hydrologic reanalyses, and the expansion of near-real-time satellite-based monitoring and analyses, particularly those describing vegetation and evapotranspiration. In the area of drought prediction, successes highlighted in the papers include the development of the North American Multimodel Ensemble (NMME) suite of seasonal model forecasts, an established basis for the importance of La Niña in drought events over the southern Great Plains, and an appreciation of the role of internal atmospheric variability related to drought events.

Despite such progress, there are still important limitations in our ability to predict various aspects of drought, including onset, duration, severity, and recovery. Critical challenges include (i) the development of objective, science-based integration approaches for merging multiple information sources; (ii) long, consistent hydrometeorological records to better characterize drought; and (iii) extending skillful precipitation forecasts beyond a 1-month lead time. Unlike other natural disasters, drought events evolve slowly in time and their impacts generally span a long period of time. Such features do make possible a more effective drought mitigation of the most adverse effects, provided a timely monitoring of an incoming drought is available.

Among the several proposed drought monitoring indices, the Standardized Precipitation Index (SPI) has found widespread application for describing and comparing droughts among different time periods and regions with different climatic conditions. However, limited efforts have been made to analyze the role of the SPI for drought forecasting. The aim of the paper is to provide two methodologies for the seasonal forecasting of SPI, under the hypothesis of uncorrelated and normally distributed monthly precipitation aggregated at various time scales k . In the first methodology, the auto-covariance matrix of SPI values is analytically derived, as a function of the statistics of the underlying monthly precipitation process, in order to compute the transition probabilities from a current drought condition to another in the future.

The proposed analytical approach appears particularly valuable from a practical stand point in light of the difficulties of applying a frequency approach due to the limited

number of transitions generally observed even on relatively long SPI records. Also, an analysis of the applicability of a Markov chain model has revealed the inadequacy of such an approach, since it leads to significant errors in the transition probability as shown in the paper.

In the second methodology, SPI forecasts at a generic time horizon M are analytically determined, in terms of conditional expectation, as a function of past values of monthly precipitation. Forecasting accuracy is estimated through an expression of the Mean Square Error, which allows one to derive confidence intervals of prediction. Validation of the derived expressions is carried out by comparing theoretical forecasts and observed SPI values by means of a moving window technique. Results seem to confirm the reliability of the proposed methodologies, which therefore can find useful application within a drought monitoring system.

1.3 AGRICULTURAL DROUGHT FORECASTING AND WARNING SYSTEM

Damage from agricultural droughts spread gradually on wide areas over a long period. In addition, due to climate change, the frequency and intensity of drought have increased. The drought of 2015 in Korea has been the worst due to lowest rainfall. It has been suggested that Korean government revisit the agricultural water management policy and incorporate in it the drought forecasting and warning system for effective drought management. In this study, new information of preparedness and mitigation planning for agricultural drought in Korea, including a long term plan and water saving strategy has been suggested. It analyses the previous drought status using the agricultural drought map, the prediction information related to drought damages, and the drought-related response capability of irrigation facilities. The system can be used to provide current and future drought conditions and the comprehensive data of drought evaluation for supporting decision makers. Also, the agricultural drought forecasting and warning system can give an opportunity to build the proper response for agricultural drought holistically and systematically.

Implementation of effective drought management policies requires both advanced technologies and appropriate methods. Monitoring and forecasting systems, practical tools for risk assessment, as well as simple and objective criteria to select and implement appropriate drought mitigation measures are key elements for a successful

drought management strategy. These key issues have been tackled by universities and public agencies involved in the EU projects Sedemed and Sedemed II (Programme Interreg IIIB MEDOCC), aimed at the definition of an integrated network for real time monitoring of drought, the development of common methodologies for drought analysis and forecasting, as well as the definition of proper mitigation strategies for the Mediterranean countries.

The book presents the main outcomes of such projects with a special focus on: drought monitoring and forecasting techniques at different spatial scales; new or modified agrometeorological indices and remote sensing technique for drought identification and characterization; tools to improve surface water resources management under drought conditions; methods and tools for groundwater resources monitoring and management, based on hydrogeological and hydrodynamics characteristics of aquifers; general criteria to select and implement mitigation strategies to prevent or minimize drought impacts.

California is the only state in the nation with no statewide groundwater regulation.

It is clear that for California to address current and future water challenges, the state must adopt a more comprehensive approach to groundwater management. California has recently taken bold steps to re-think its approach to groundwater. Yet, a commitment to sustainable groundwater management is only as good as the metrics that will be used to assess progress. Further, climate change is already altering the timing and reliability of California's water supplies. Groundwater management must incorporate the state's changing hydrology and its impact on groundwater recharge and use. This paper reviews a number of approaches to groundwater management. California's groundwater policy can benefit from the lessons learned elsewhere. Specifically, California's groundwater policy should provide clear metrics to define and measure sustainable yield, explicitly incorporating climate change.

CHAPTER 2

AIM AND SCOPE

2.1 AIM OF PROJECT

The main aim of this project is, Analysing and Predicting the drought based on humidity, temperature, ground water level, rainfall, wind speed. This motivates the integration of both methods to combine drought forecasts from different sources for improving drought prediction. Meanwhile, a variety of important contributing factors, including local climate, evaporation changes, and human activities may maintain or amplify different types of drought. Seasonal drought prediction mainly focuses on fully utilizing these predictability sources or contributing factors operating at local, continental, or global scales to facilitate drought prediction at seasonal time scales.

2.1.1 OBJECTIVES

The main objective of this project is, Analysing and Predicting the drought based on humidity, temperature, ground water level, rainfall, wind speed.

2.2 SCOPES

Unlike other natural disasters, drought events evolve slowly in time and their impacts generally span a long period of time. Such features do make possible a more effective drought mitigation of the most adverse effects, provided a timely monitoring of an incoming drought is available. Among the several proposed drought monitoring indices, the Standardized Precipitation Index (SPI) has found widespread application for describing and comparing droughts among different time periods and regions with different climatic conditions. However, limited efforts have been made to analyze the role of the SPI for drought forecasting. The aim of the paper is to provide two methodologies for the seasonal forecasting of SPI, under the hypothesis of uncorrelated and normally distributed monthly precipitation aggregated at various time scales k .

In the first methodology, the auto-covariance matrix of SPI values is analytically derived, as a function of the statistics of the underlying monthly precipitation process, in order to compute the transition probabilities from a current drought condition to another in the future.

The proposed analytical approach appears particularly valuable from a practical stand point in light of the difficulties of applying a frequency approach due to the limited number of transitions generally observed even on relatively long SPI records.

Also, an analysis of the applicability of a Markov chain model has revealed the inadequacy of such an approach, since it leads to significant errors in the transition probability as shown in the paper. In the second methodology, SPI forecasts at a generic time horizon M are analytically determined, in terms of conditional expectation, as a function of past values of monthly precipitation. Forecasting accuracy is estimated through an expression of the Mean Square Error, which allows one to derive confidence intervals of prediction. Validation of the derived expressions is carried out by comparing theoretical forecasts and observed SPI values by means of a moving window technique. Results seem to confirm the reliability of the proposed methodologies, which therefore can find useful application within a drought monitoring system.

The ultimate objective of drought prediction is to prepare a mitigation plan in advance, rather than resolve intellectual curiosity about nature. Drought forecasting plays an important role in mitigating the negative effects of drought; hence, various approaches for predicting droughts have constantly been attempted, such as stochastic methods, combined statistical and dynamical models, categorical prediction, machine learning approaches, and hybrid models. However, drought prediction is still challenging because, in addition to precipitation deficit, complicated interactions among other variables, such as temperature, evapotranspiration, land surface processes, and human activities, also contribute to droughts.

Further, meteorological anomalies, due to climate changes have made it increasingly difficult to predict precipitation, which forms the basis for forecasting drought. Agricultural droughts determined by soil moisture must be predicted several months ahead for proper and rapid resource allocation, because this allocation can mitigate the effects of upcoming droughts by supplying timely water and guaranteeing suitable crop growth and availability of food Water doi:10.3390/w11040705 www.mdpi.com/journal/water Water 2019, 11 of 15 resources. Droughts are generally classified into four categories, namely, meteorological, hydrological, agricultural, and socio-economic droughts. In recent years, the prevalence of machine learning methodologies, and frequent droughts and floods

around the world, have increased the prediction of agricultural drought.

However, the uncertainties of prediction caused by combination of meteorological factors, still remain a problem.

According to the Intergovernmental Panel on Climate Change (IPCC) AR5 guidance note, the complex use of different models, complexity of models, and inclusion of additional processes in the analysis are the main reasons for the increase in uncertainties. Thus, the complex models used for the integrated analysis of meteorological and agricultural drought have increased the uncertainty in the results. However, agricultural drought prediction methods that do not include spatial information, and are only based on precipitation, such those put forth in cannot predict the spatial distribution of agricultural drought. Another difficulty in predicting agricultural drought stems from predictions based on the meteorological pattern, such as patterns of past precipitation.

Even well-designed agricultural drought models, based on meteorological factors, cannot make accurate predictions because of the shifts in existing patterns, due to climate change. According to a study of agricultural forecasting models from 2007 to 2017 precipitation was used as an input variable in almost all the models. Previous studies have attempted to develop drought prediction models by using remote sensing data. For example, Sheffield et al applied a downscaling technique to predict seasonal droughts by combining hydrological and agricultural models. However, such prediction models are difficult to use because they require decades of weather data. Hao et al. designed a system that predicts soil moisture and severe drought affected areas by focusing on seasonal droughts, which is similar to the approach used in our model.

However, because of its global scale and large system, it is difficult to use for developing regional drought mitigation plans. Therefore, alternative agricultural drought-forecasting methods that are easy to use and scalable are required to realize practical drought mitigation plans. Moreover, these methods must consider the uncertainty of the precipitation forecasts and spatial information. To achieve this objective herein, we designed a model, called the severe drought area prediction (SDAP) model, to estimate soil moisture index (SMI) maps after several months using surface factors. All variables were composed of surface factors derived from remote sensing data, which are related to soil moisture, to obtain spatial information of agricultural drought.

The SDAP model predicts the area of severe agricultural drought in terms of preparation information to help mitigate agricultural drought in case there exists a possibility of meteorological drought. This model was proposed to help planning of priority and rapid water allocation for predominantly drought-affected risk area in occurring a drought. Therefore, prediction, in the SDAP model, does not imply a prediction of the occurrence of drought. Rather, it provides information on the future development and evolution of agricultural drought, assuming that rainfall-deficit conditions continue to prevail, in preparation for drought mitigation plans.

We believe this is a realistic approach that avoids the uncertainty problem associated with meteorological drought forecasting. In order to design the SDAP model, we classified the land surface factors that affect the soil moisture into four categories, which are vegetation, topographic, thermal, and water factors during the non-rainfall period. Thermal increase is one of the core factors that increases the risk of agricultural drought by causing loss of soil moisture due to increased evaporation, while the vegetation factor delays the loss of soil moisture by slowing the increase in surface heat. The topography factor is another important determinant of soil moisture.

The water content status is also a factor related to the soil moisture remaining after the drought period. These environmental factors, which are used as the initial land conditions as a predictor for drought forecasting, are regressed on the soil moisture after the non-rainfall period and can then predict soil moisture during short-term drought. In this regard, our model is designed to regress 15 input variables (derived from four surface factors) to the soil moisture index using the random forest (RF) algorithm.

Water of 15 The RF algorithm, proposed by Breiman , is an ensemble technique that uses an average of the multiple decision trees to reduce overfitting and lead to sound regression performance. Furthermore, the RF algorithm is advantageous for the analysis of large datasets; thus, it is suitable for analysis involving satellite images. It generates a tree-based regression function; hence, the scale adjustment between different features is not necessary, which is a considerable advantage in cases involving multiple features with different scales, as is the case in the current study.

The novelty of this study is as follows: First, we used simple data. We derived the drought function from the selected training area in the years when drought existed.

In contrast, other studies trained with the entire study area or used data for years. Second, once created, the drought function was repeatedly used for the same season and local conditions.

This drought function was based on the regression information for soil moisture according to the land surface and climatic conditions in the area during a non-rainfall period. Third, the precipitation data were not used because rainfall-deficit was assumed. Fourth, we used topographical data, which were an important factor for soil moisture at the local scale, as a variable.

The agricultural drought forecasting models that were developed were mostly analyzed on the global scale and mainly used satellite imagery. In contrast, our SDAP model uses terrain data for satellite imagery to improve the model to fit the local scale. To design the SDAP model, the following approach was used. We first defined four surface factors that affect the maintenance and decrease of the SMI during non-rainfall periods. We then selected input variables corresponding to this category of surface factors. We generated drought function with the RF algorithm using input variables and three months afterward SMI to train agricultural drought. Additionally, we have identified the order of feature importance that affects drought training (regression) among the input variables (features). The drought function $f(x)$ used to predict the SMI of the study area. We verified the training and prediction performance of the drought function using the actual SMI value.

2. Methods Study and Training Area Our study area covered 12 administrative district.

CHAPTER 3

SYSTEM DESIGN & METHODOLOGY

3.1 PROPOSED SYSTEM

The proposed system assess the parameters contributing to drought and predicts the probability of drought occurrence. If the occurrence of the drought is detected, then the severity of the drought is determined. This is done by analysing the historical values of the causative parameters and predicting the upcoming values. Machine learning algorithm is used for analytic purpose. Thus the predicted values is assessed to determine the drought probability.

3.2 EXISTING SYSTEM

The existing system for drought forecasting uses many techniques like data mining, machine learning and deep learning. The technique of data mining is introduced to find associations between drought and several oceanic and climatic indices and thus predicting the drought. Machine learning methods have also been used recently to forecast droughts, using drought indicators such as the Standardized precipitation Index (SPI) as data input.

Machine learning techniques like artificial neural networks (ANNs), support vector regression (SVR), and coupled wavelet-ANNs, pre-process input data using wavelet analysis and the forecast results of all three methods were compared using two performance measures. Drought can be forecasted by using machine learning based on remote sensing and long range forecasting data. Another method is Deep Belief Network consisting of two Restricted Boltzmann Machines for long-term drought prediction using lagged values of Standardized Streamflow Index (SSI) as inputs.

This method is applied to predict different time scale drought indices. Individual assessment of each and every causative parameters is not involved in the existing methods. There is a need for a tool to predict the water scarcity in an area, sometimes as small as size of a range of 1000 ha.

Most of the predictions primarily depends on precipitation rate .Precipitation rate is a major factor in causing drought but when predicting for a localized area there are certain other factors which can also have a potential impact on drought.

3.3 SYSTEMCAPABILITIES

The proposed system assess the parameters contributing to drought and predicts the probability of drought occurrence.If the occurrence of the drought is detected, then the severity of the drought is determined.This is done by analysing the historical values of the causative parameters and predicting the upcoming values.Machine learning algorithm is used for analytic purpose.Thus the predicted values is assessed to determine the drought probability.

3.4 REQUIREMENTSPECIFICATION

This proposed software runs effectively on a computing system that has the minimum requirements. Undertaking all the equipment necessities are not satisfied but rather exist in their systems administration between the customer"s machines already. So the main need is to introduce appropriate equipment for the product. The requirements are split into two categories, namely:

Software Requirements

The basic software requirements to run the program are:

- Operating System - Windows XP
- Coding language - Java
- IDE - Net beans
- Database -MYSQL

Hardware Requirements

The basic hardware required to run the program are:

- System - Pentium-IV

- Speed - 2.4GHZ
- Hard disk - 40GB
- Monitor - 15VGA color

3.5 PREDICTING OF DROUGHT AREAS

Information on drought severity is the most practical and important information for mitigating an upcoming drought. The objective of this study was to predict the distribution of soil moisture assuming non-rainfall conditions. According to the National Drought Mitigation Center (NDMC), it is necessary to make a plan for drought before beginning, confirm the priority of water dependence on agriculture and the community, and to immediately apply this step-by-step when water begins to become insufficient. In this step, the SDAP model can help by supporting spatial information on droughts in advance, and allow effective and planned resource allocation to reduce the impact of upcoming droughts.

The SDAP model can show future severe agricultural drought areas through non-rainfall periods because the method can reduce the uncertainty by separating two models (i.e., meteorological and agricultural droughts) with different features in the analytical results. Under non-rainfall conditions, the SDAP model works on the assumption that meteorological drought forecasting is preceded. Thus, the SDAP model can analyze drought without considering meteorological data. However, the absence of meteorological data for analysis does not mean that we have excluded climate factors.

Machine learning on the growth of soil drought is a learning of the trend caused by climate and seasonal changes in the area. Therefore, the further away from the training area, the more the climate and the weather become different; hence, the prediction error of the SMI increases. We found that one of the studies predicted soil moisture using hybrid machine learning without climate data, with a similar concept to the SDAP model.

However, that study cannot show the spatial distribution of the SMI because of the usage of the actual measured data from the soil of only seven points, not the remote sensing data. In fact, the prediction by aimed to only predict the soil moisture of seven trained points using training data (corresponding to the validation of the performance of Water 2019, 11, 705 11 of 15 the drought function in our study).

Hence, this is just an assessment of the training performance, and not prediction. However, in our study, the training data were only used to generate the drought function, and different data from another year were used for prediction. Another study, using only meteorological data, cannot be used in practical drought mitigation plans because of the exclusion of the spatial distribution of agricultural drought in the prediction results. In contrast, the results of the SDAP model can show clear SMI trends including the severe area of the SMI .

The mitigation of Short-term drought requires the SMI spatial trend information rather than accurate SMI values . In this model, although the SMI error (RMSE = 0.382, MAE = 0.375) of the predicted study area increased compared to the training performance of the drought function (RMSE = 0.052, MAE = 0.039), our model has achieved the aim sufficiently because it can show the agricultural drought trend clearly Feature Importance We found that the thermal factor is one of the major factors affecting soil moisture during a drought period by confirming the findings of Sruthi and Aslam who noted that increased temperature reduces soil moisture.

Furthermore, TIRS1 in the thermal factor was the most important feature in the drought function. However, one should be cautious when considering the feature importance. For example, relative vegetation features have been shown to be of low importance in our model. However, according to Sruthi and Aslam, vegetation and temperature were strongly correlated with each other; hence, vegetation might play an important role in preserving soil moisture. Therefore, feature importance should be carefully considered based on other studies. We obtained this variable importance information because we used the RF algorithm for drought training. The RF algorithm has the advantage of providing variable importance; hence, it is also used to extract important variables and remove unnecessary variables when creating a model.

One case involved the estimation of the relative importance of variables related to soil hydrological properties in the field of hydrology using this importance function. However, whether the correlation between the SMI and feature importance is negative or positive, is unknown.

Therefore, it is also helpful to reference the multi-linear regression coefficients to understand feature importance. Comparing the feature importance of the drought functions, calculated from various regions and periods, can provide important insights into drought studies related to soil moisture.

Limitations and Applications of the SDAP Model The predicted SMI, using machine learning applied to satellite images with the resolution of 30 m, is a suitable drought information for planning the mitigation of short-term drought at the regional level because the local spatial distribution is acquired. However, some general remote sensing limitations exist, such as difficulty in obtaining data of the right date to be analyzed because of cloud or revisit cycle of the Landsat-8. This limitation can be addressed by considering using other satellite images with a higher temporal resolution. This solution remains to be confirmed in further studies.

Along with the general limitations of remote sensing, this SDAP model has the following limitations: First, the SDAP model applies only to the prediction of short-term droughts, within several months, because the SMI is a suitable measure for that timescale. In practice, the soil will become dry, regardless of the surface state if non-rainfall conditions persist for longer periods. Therefore, the SDAP model is suitable for short-term droughts that can be mitigated by human action within a few months. Second, the prediction of droughts may be difficult if the area has no drought experience or different seasons, because this model is based on learning from past droughts. Therefore, various drought functions by period or season must be considered to make a suitable prediction of the situation. Third, this model must be based on a sequential approach to finding solutions by analyzing droughts by each expert group. Thus, it is not suitable for an integrated method of calculating Water 2019, 11, 705 12 of 15 meteorological, agricultural, and hydrological droughts at one time because this model predicts the trend of agricultural drought change after the meteorological drought analysis is preceded.

For example, the spatial information of agricultural drought, which is the result of this model, is then subject to a primary review against the corresponding hydrological drought. After all the analyses have been conducted, and the areas of concern droughts have been identified, priority plans for water reserves (e.g., water demand control) and water allocation for these areas should be established. Therefore, the proposed model is

only suitable to prepare for short-term droughts and find focalized solutions by each expert group, instead of performing an integrated analysis of various models with high uncertainty.

5. Conclusions Agricultural drought is a disaster that must be managed effectively as it directly affects food security.

For this reason, several models have been developed to predict agricultural drought. However, regardless of how good a model is, inaccuracies arise in the prediction of droughts, due to meteorological anomalies where the model is based on repeated patterns or historical precipitation data. The proposed SDAP model was hence developed by accepting those uncertainties, rather than overcome meteorological uncertainties. This SDAP model does not depend on precipitation data because it predicts potential drought-severe areas under the assumption that rainfall-deficit conditions already exist.

In addition, existing other models can be catastrophic in the event that an unpredicted drought occurs. However, given that it makes prediction under the condition of no rainfall such as that done in our model, there is a benefit for society even if no drought occurs. Although our model is designed for a short-term drought, accumulation of these predictions also helps us identify areas that are susceptible to drought so that mitigation plans can be prepared from a long-term perspective. Given that the ultimate objective for predicting droughts is to develop mitigation plans, the model must be a simple, practical, and useful such as the SDAP model. This model is easy to understand as it uses a regression algorithm compared to other huge prediction systems. Further, the analysis of various case studies for many regions and drought functions in SDAP model can provide meaningful insights into the key factors associated with drought.

At present, we are working on a comparison study of machine learning algorithms to improve the SDAP model as well as a case study involving the application of the SDAP model. Agricultural drought is a disaster that must be managed effectively as it directly affects food security. For this reason, several models have been developed to predict agricultural drought. However, regardless of how good a model is, inaccuracies arise in the prediction of droughts, due to meteorological anomalies where the model is based on repeated patterns or historical precipitation data. The proposed SDAP model was

hence developed by accepting those uncertainties, rather than overcome meteorological uncertainties. This SDAP model does not depend on precipitation data because it predicts potential drought-severe areas under the assumption that rainfall-deficit conditions already exist. In addition, existing other models can be catastrophic in the event that an unpredicted drought occurs.

However, given that it makes prediction under the condition of no rainfall such as that done in our model, there is a benefit for society even if no drought occurs. Although our model is designed for a short-term drought, accumulation of these predictions also helps us identify areas that are susceptible to drought so that mitigation plans can be prepared from a long-term perspective. Given that the ultimate objective for predicting droughts is to develop mitigation plans, the model must be a simple, practical, and useful such as the SDAP model. This model is easy to understand as it uses a regression algorithm compared to other huge prediction systems. Further, the analysis of various case studies for many regions and drought functions in SDAP model can provide meaningful insights into the key factors associated with drought. At present, we are working on a comparison study of machine learning algorithms to improve the SDAP model as well as a case study involving the application of the SDAP model.

3.6 DROUGHT PREDICTED IN DIFFERENT AREAS

Drought is a natural disaster of below-average precipitation in a certain area caused by disruption to an expected precipitation pattern, and it has a very wide impact. One type of drought that occurs in Indonesia is meteorological drought, which can be defined as an event that takes place naturally and repeatedly because of reduced rainfall from normal conditions. One of the provinces in Indonesia with the most frequent meteorological droughts is East Nusa Tenggara (NTT), and this is listed as the top priority region that is most vulnerable to drought. Mitigation of drought, by methods such as providing reliable predictions of future droughts, is definitely required to minimize the risk and the negative impact of drought.

Meteorological drought monitoring can be done by examining the rainfall data recorded from observation stations; such data are effective and relatively accurate in describing the

rainfall in an area. However, meteorological stations are not evenly distributed, resulting in the reduced accuracy of the results of the analysis particularly in regions with no stations. To deal with this drawback, a dataset of satellite-based data has been extensively used in many analyses. Remote-sensing data originating from meteorological satellites can provide rainfall data with a better spatial and temporal distribution. According to Rhee and Im remote-sensing data can be used widely and dynamically in drought monitoring. For tropical regions, the remote-sensing data generated from meteorological satellites that have been widely used in various studies are from the Tropical Rainfall Measuring Mission. Most studies have found that TRMM satellite precipitation gives a relatively good performance. For the case of Indonesia, the performance of the TRMM has been investigated using TRMM data for drought analysis.

Another dataset that has been extensively used to build drought prediction is Modern-Era Retrospective Analysis for Research and Applications. This is a re-analysis product that assimilates satellite rainfall estimates from the Special Sensor Microwave Imager and the TRMM Microwave Imager. Among the studies that have used to build drought prediction are that of Kulkarni, who applied to the case of India and who compared the performance of with that of other re-analysis products such as Interim and the re-analysis for China, and found that has a better performance has been used in numerous studies in different climatic regions. Uncertainties in datasets have been evaluated against different observations. The results showed that provides valuable information consistent with observation, especially in the mid-latitudes, while uncertainties in the high latitudes are often large.

This present paper investigates the performance of TRMM and for predicting drought in East Nusa Tenggara, Indonesia. The prediction of drought will generate a classification of drought, based on the 3-month SPI derived from those two data sources. The 3-month SPI is used because it can describe short- and medium-term humidity conditions, according to the World Meteorological Organization (WMO). In addition, the 3-month SPI has commonly been used by the Indonesian Agency for Meteorology, Climatology and Geophysics for monitoring drought conditions in Indonesia. Following Rhee drought is predicted using several predictors such as the Normalized Difference Vegetation Index (NDVI), the average soil surface temperature day and night, the Multivariate ENSO Index (MEI), and the Arctic Oscillation Index (AOI).

This research applies two different machine learning methods to classify the drought status, Classification and Regression Tree (CART) and Random Forest (RF). Both methods were selected because of their strength in applications to a large sample dataset, as in our case. Moreover, both methods have been proved to be computationally efficient. Various machine learning approaches have been extensively applied in the case of drought prediction . The most recent work by Fung et al provides a comprehensive review of the applications of statistics-based modelling as well as machine learning methods for drought forecasting over the period from 2007 to 2017. Most of the papers agree that machine learning is a powerful tool for drought forecasting. This present paper also proposes the combination of the machine learning methods with sampling method to overcome the problem of imbalance class response as well as to improve the predictive performance.

Drought prediction is of critical importance to early warning for drought managements. This work provides a synthesis of drought prediction based on statistical, dynamical, and hybrid methods. Statistical drought prediction is achieved by modeling the relationship between drought indices of interest and a suite of potential predictors, including large-scale climate indices, local climate variables, and land initial conditions. Dynamical meteorological drought prediction relies on seasonal climate forecast from general circulation models (GCMs), which can be employed to drive hydrological models for agricultural and hydrological drought prediction with the predictability determined by both climates forcing and initial conditions.

SVM model has been applied for classification of real time data obtained from Meteorological department. Here we considered parameters like maximum rainfall, minimum rainfall and precipitation. The prediction is base on the dataset collected for a period of ten years. Challenges still exist in drought prediction at long lead time and under a changing environment resulting from natural and anthropogenic factors. The evaluation and forecasting of drought is made possible by the use of drought indices. Application like drought prediction depends on the measurement of the environmental factors that are based on recorded data. Severity of drought prediction using the traditional methods is complicated by the variant phenomenon of natural parameters. Efficient drought prediction is not possible in traditional methods as they lack the real time information.

One of the toughest issues to resolve in deep learning has nothing to do with neural networks; it's the matter of obtaining the correct information within the right format. Getting the correct information suggests that gathering or distinguishing the information that correlates with the outcomes you would like to International Journal of Advances in Computer Science and Cloud Computing in Drought Prediction using Machine Learning Algorithm 2 predict; i.e. information that contains an indication regarding events you care regarding. the information must be aligned with the matter you're trying to resolve. Kitten footage don't seem to be terribly helpful once you're building a facial identification system. validating that the information is aligned with the matter you ask for to resolve should be done by a knowledge human.

If you are doing not have the correct information, then your efforts to create AI resolution which should come back to the information assortment stage. The right finish format for deep learning is mostly a tensor, or a multi-dimensional array. therefore, information pipelines engineered for deep learning can typically convert all information – be it pictures, video, sound, voice, text or statistic – into vectors and tensors to that algebra operations are often applied. That information oft must be normalized, standardized and clean to extend its utility, and people where all steps in machinelearning ETL. Deep learning offers the ETL tool to perform those information preprocessing tasks.

Deep learning, and machine learning additional typically, wants an honest coaching data set to figure properly. Assembling and constructing the coaching set – a large body of known information – takes time and domain-specific data of wherever and the way to collect relevant info. The coaching set acts because the benchmark against that deeplearning nets are trained. Hence, they learn to reconstruct before they're unleashed on information they haven't seen before. At this stage, knowledgeable humans got to notice the correct data and remodel it into a numerical illustration that the deep-learning algorithmic program will perceive, a tensor. Building a coaching set is, in a sense, pre-pre-training. Machine learning generally works with 2 data sets: training-set and test set. All three ought to at random sample a bigger body of data. The first set of data is that the network a neural network. The second set is your test set. It functions as a seal of approval, and you don't use it till the top. when you've trained and optimized your knowledge, you check your neural internet against this final sampling.

The results it produces ought to validate that your internet accurately recognizes images or acknowledges them at least 'x' percentage of them. If accurate predictions are not obtained, then the coaching set must be retrained. Here we investigate the hyper parameters accustomed to tune the network. This process enhances the quality of knowledge the pre-processing technique. Indroughtseverityprediction data cleaning is the initial step that has been done by collected the required real time dataset which contains the parametersfor predicting drought using some of the machine learning algorithm.

Data preprocessing is a very important and quite underestimated step in Machine Learning pipelines. It provides cleaned and relevant datasets which then can be used in further steps like classification or regression. Thedata which is fed to the SVM classifier to predict if a given image segment belongs to foreground or background. After collection of measurement data such as maximum temperature, minimum temperature, precipitation and rainfall, data cleaning is done to remove irrelevant data, incomplete data, missing data so that it can be fed into the SVM which gives the result if the drought has occurred. Data collection is done which contains the parameters required for predicting drought like minimum temperature, maximum temperature, precipitation and rainfall. After that the incomplete data, inaccurate data, inconsistent data, missing data, duplicate data are all resolved by using data cleaning.

Data cleaning is done for the data which is irrelevant and the data which is more frequentisused.Support Vector Machine (SVM): In machine learning, support-vector machines are supervised studying models with associated gaining knowledge of algorithms that analyze facts used for classification and regression analysis. Given a fixed of education examples, every marked as belonging to one or the other of two classes, an SVM education algorithm builds a version that assigns new examples to 1 category or the other, making it a nonprobabilistic binary linear classifier (although strategies inclusive of Platt scaling exist to apply SVM in a probabilistic category setting).

An SVM model is an illustration of the examples as factors in space, mapped so that the examples of the separate categories are divided through a clear gap which is as wide as possible.New examples are then mapped into that identical area and predicted to belong to a category based totally on which aspect of the space they fall.

Further to acting as linear category, SVMs can efficaciously carry out a non-linear classification using what's known as the kernel trick, implicitly mapping their inputs into excessive-dimensional function spaces. Drought has a great impact on agriculture, economy not only in India but across the whole world. In this paper we have proposed a method for drought prediction after analysis of dataset.

This system describes the drought prediction scheme using the support vector machine algorithm. Frontier technologies and approaches are used to solve the practical problem of adverse natural phenomenon in large area. Developing applications using this can correlate and predict natural calamities ranging from earthquakes to hurricanes that are normally occurring in the normal fashion but may have cross correlation to observable physical parametric variations around the globe. This will be of great advantage.

3.7 MODULES

A module is a building block for creating experiments. Each module encapsulates a specific machine learning algorithm, function, or code library that can act on data in your workspace. The modules are designed to accept connections from other modules, to share and modify data.

Here four modules are used. They are as follows

3.7.1. Analysis Module

3.7.2. Drought Prediction Module

3.7.3. Calculate Severity index

3.7.4. View Information Module

3.7.1 Analysis Module

In this Module, analysing the historical values of the causative parameters and predicting the upcoming values. Machine learning algorithm is used for analytic purpose.

3.7.2 Drought Prediction Module

This method predicts drought by assessing impact of individual factors on drought, analyzing the relationship in them and draws the conclusion regarding the drought.

Most of the predictions primarily depends on precipitation rate, and basic parameters like humidity, temperature, ground water level, rainfall, wind speed for predicting drought. This is also a less complex method in predicting drought comparing with existing drought predicting methodologies. drought prediction problem using support vector machine algorithm.

Data is collected for large period and stored in a file. In our application data collected for several years is used for drought prediction. We proposed a support vector machine algorithm for long term drought prediction. The drought occurs mainly due to lack of precipitation and due to high temperature. Support vector machine model was found to provide better prediction results, recording lower prediction errors and therefore can be more reliable and efficient for long term drought prediction

3.7.3 Calculate Severity Index Module

In this Module, Calculate the severity index based on drought prediction parameters like humidity, temperature, ground water level, rainfall, wind speed for predicting drought.

3.7.4 View Information Module

In this Module, user can view the information about whether the drought will be occurred or not. It will be displayed on the screen as drought prone and no drought.

CHAPTER 4

SOFTWARE DEVELOPMENT METHODOLOGY

4.1 DESCRIPTION OF DIAGRAM

In the Fig4.1, It is the block diagram of the drought prediction. If we see the top of the diagram there is data collection. At first the data is collected from the website called kaggle websites. Kaggle website is a data search platform. Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers. And next the data preparation is the process of collecting, cleaning and consolidating data into one file or data table. Here we have prepared the data in the form of table. Data retrieval obtains the data from the database. In order to retrieve the desired data the user presents a set of criteria by query. After preparing and retrieving the data we apply the machine learning algorithm called Random Forest algorithm. With the help of this algorithm we can predict the drought. And it will display about the drought.

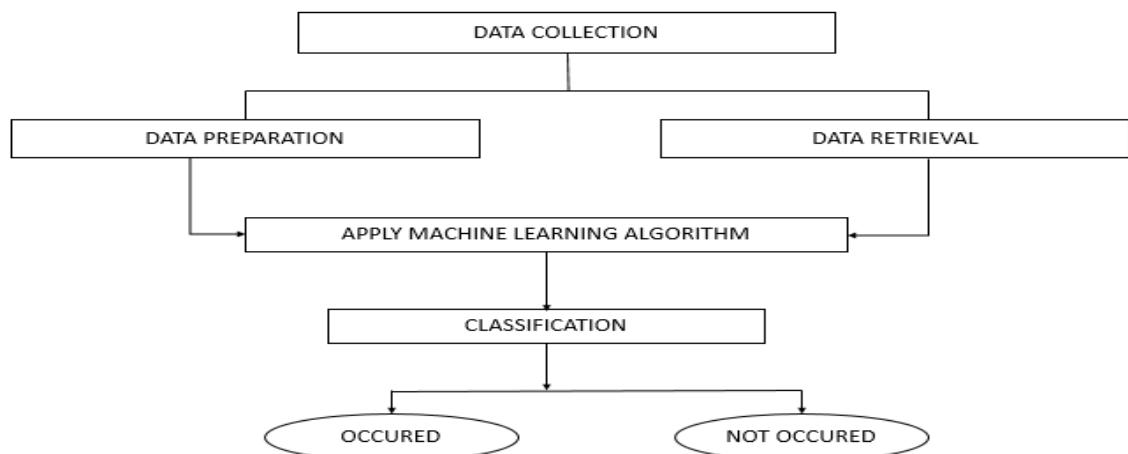


Fig 4.1: Block Diagram

4.2 ARCHITECTURAL DESIGN

The above figure is the Architecture diagram of the drought prediction app. By following the above Architecture diagram app is created. First of all we have to collect the datasets in order to predict the drought. These datasets are available in kaggle website. We can get all kinds of information there. Then in the data sets we have to retrieve the data from the obtained data. Then the obtained data will be undergoing the data preparation in that process it undergoes data processing where and data wrangling is also will be done. Then the feature will be extracted from the data sets and the feature scaling is also done by feature scaling and selection. Here machine learning algorithm used here is random forest. By applying this algorithm we can do modeling and we can also evaluate and also perform the tuning. And then finally deployment and monitoring is done.

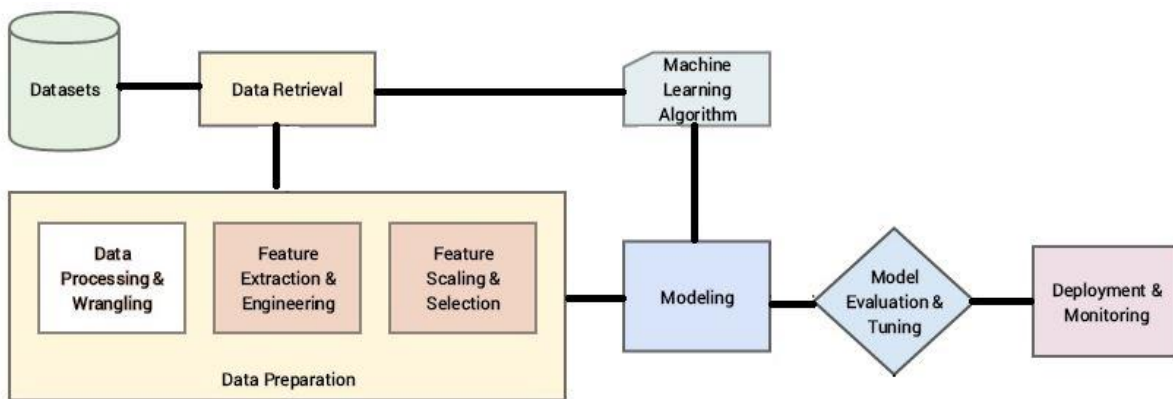


Fig 4.2: Architectural Diagram

4.3 SEQUENCE DIAGRAM

Sequence diagram model the flow of logic within your system in a visual manner, enabling you both to document and validate your logic, and commonly used for both analysis and design purpose. Sequence diagram are the most popular UML artifact for dynamic modeling, which focuses on identifying the behavior within your system.

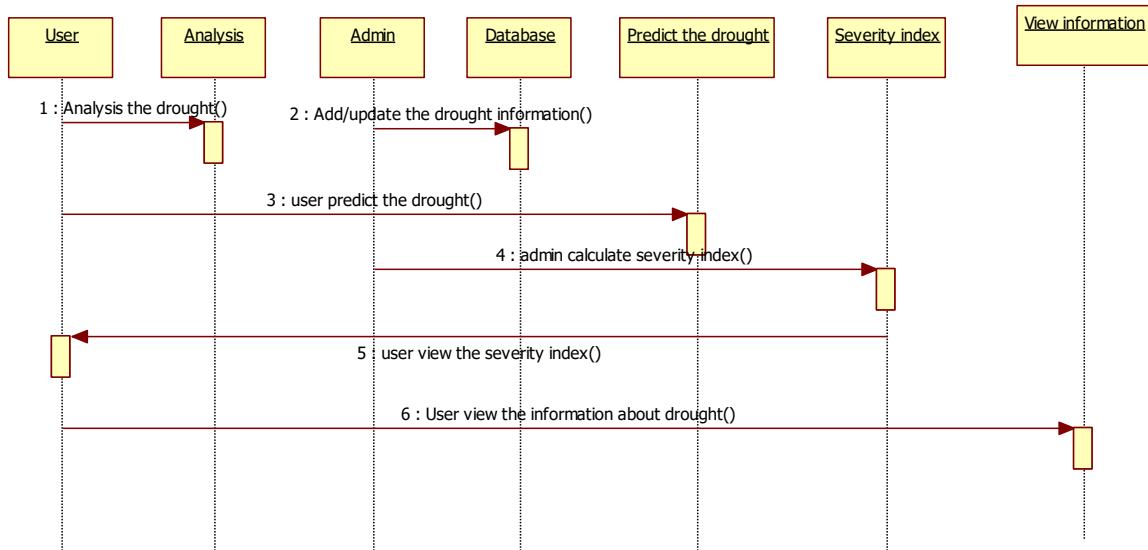


Fig 4.3:Sequence Diagram

4.4 USE CASE DIAGRAM

Use case diagrams overview the usage requirement for system. They are useful for presentations to management and/or project stakeholders, but for actual development you will find that use cases provide significantly more value because they describe “the meant” of the actual requirements. A use case describes a sequence of action that provides something of measurable value to an action and is drawn as a horizontal ellipse.

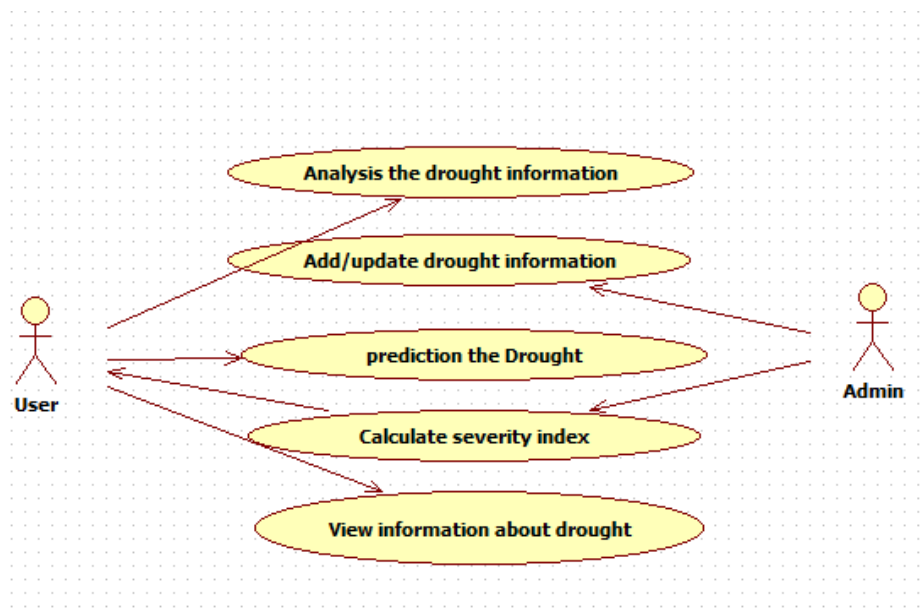


Fig 4.4:Use case Diagram

4.5 COLLABORATION DIAGRAM

Another type of interaction diagram is the collaboration diagram. A collaboration diagram represents a collaboration, which is a set of objects related in a particular context, and interaction, which is a set of messages exchange among the objects within the collaboration to achieve a desired outcome.

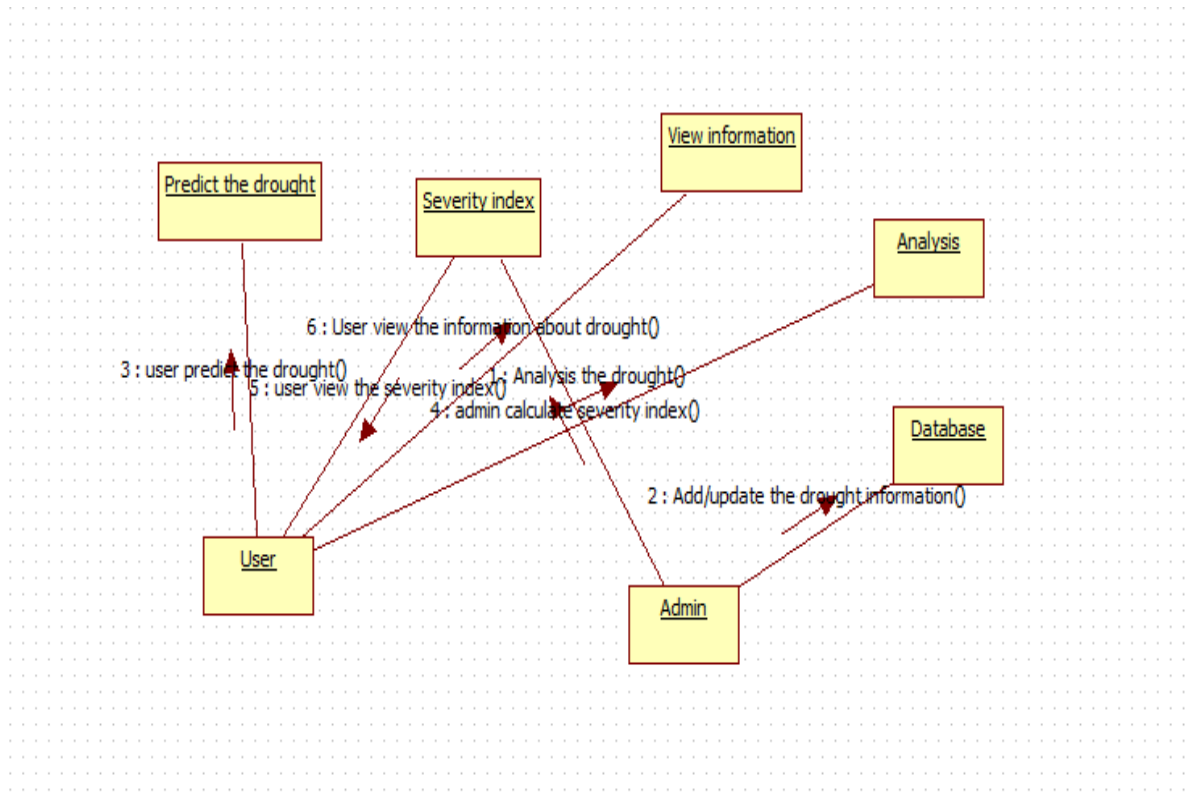


Fig 4.5:Collaboration Diagram

4.6 ACTIVITY DIAGRAM

Activity diagram are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. The activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. Activity diagram consist of Initial node, activity final node and activities in between.

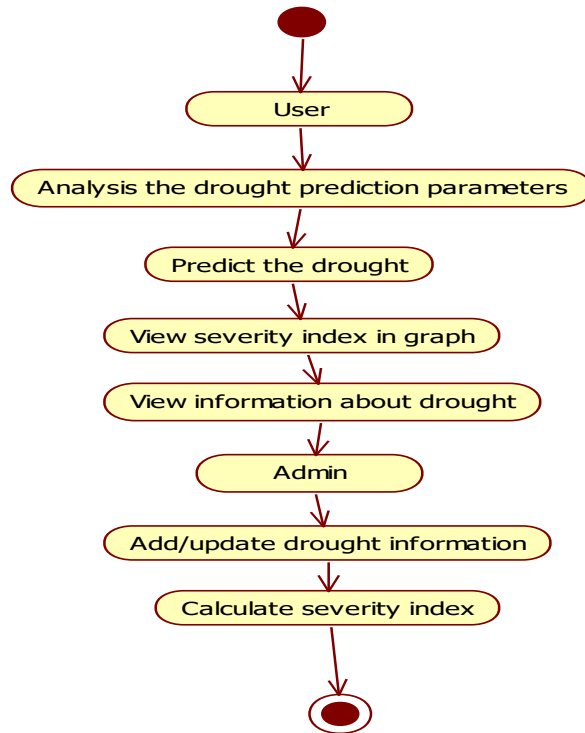


Fig 4.6:Activity Diagram

CHAPTER 5

RESULTS AND DISCUSSION

This method predicts drought by assessing impact of individual factors on drought,analyze the relationship in them and draws the conclusion regarding the drought.Analysing and Predicting the drought based on humidity, temperature, ground water level, rainfall, wind speed.

The Fig 5.1 is the fig where the code is located in order to run the program we have to copy the path and run it in the anaconda prompt.All of the algorithms and the code is located here.Here I have created a folder a file named Project and I have inputted all of my inputs and the code.

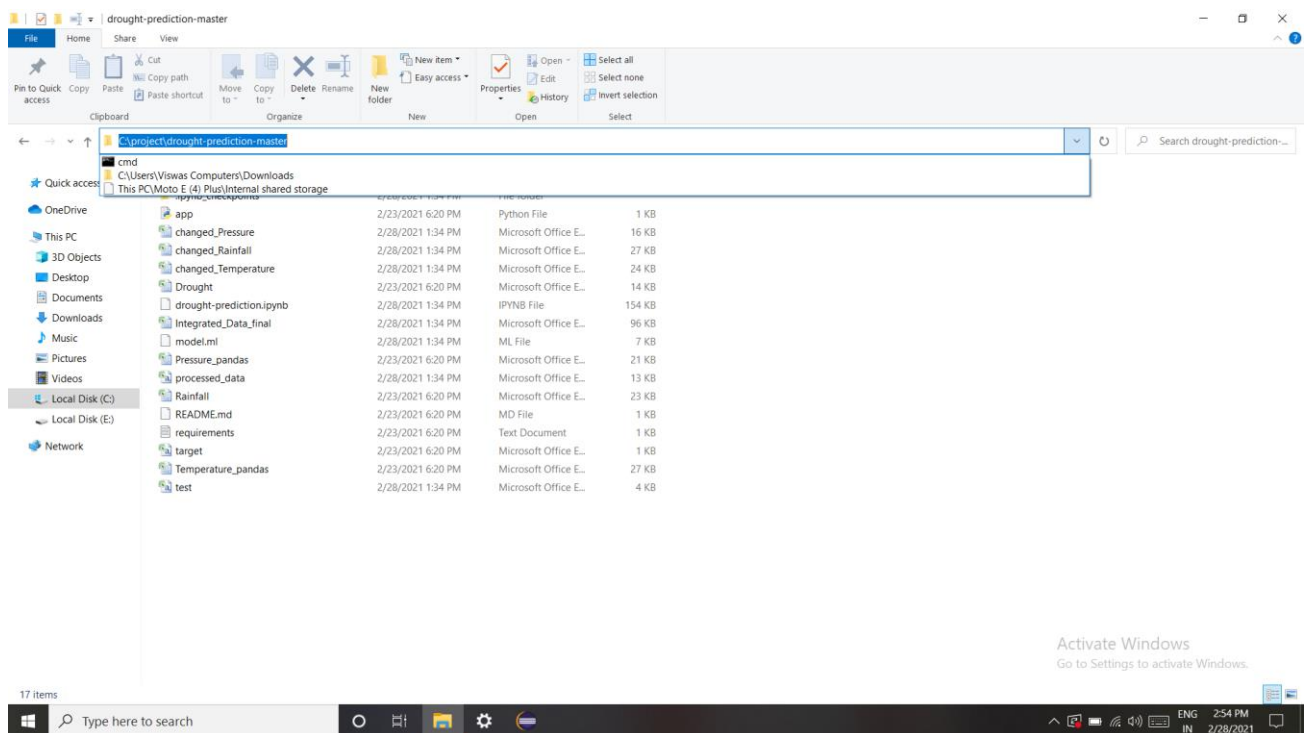


Fig 5.1:To Run the Program

The Fig 5.2 is the anaconda command prompt.This debugs the code and runs the app.By running this command prompt the port will start running.Without running this one we cannot start or run the app.In order to run this network connectivity should be very speed.Based on the internet connectivity and the speed it runs.But this consumes very less net.This plays a major role in the running of the code.This helps the code to run faster with consuming less time.If we see in the below figure after copy and pasting the link that contains the code and the inputs.Then the debugging mode will be on ON mode.And then it restarts with stat.By this the Debugger will become active.The debugger pin will be dispalyed like below.Then the code will start to run on the port that is as shown below.Now we can close the window or we can just minimize it.And then we can proceed to the next step.

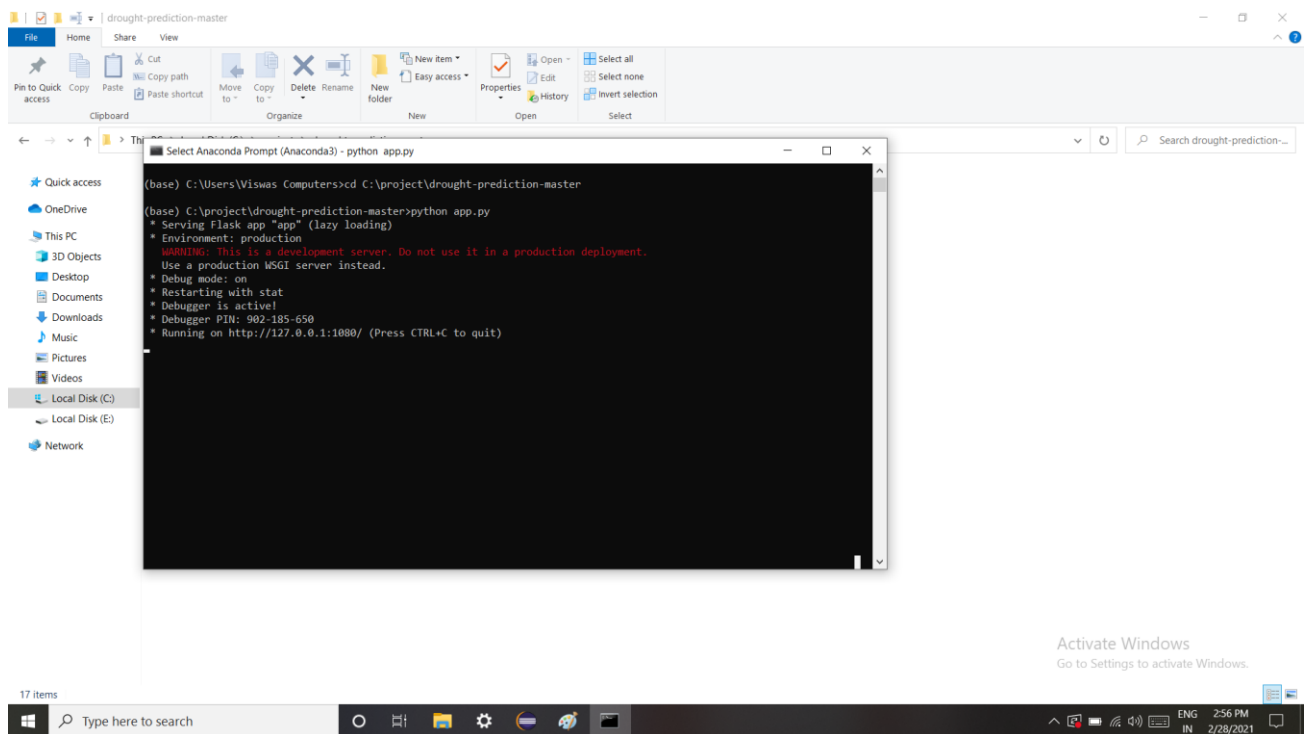


Fig 5.2:Anaconda Command Prompt

The screenshot that is shown below is eclipse ide workspace. Eclipse is an integrated development environment (IDE) used in computer programming. It contains a base workspace and an extensible plug-in system for customizing the environment. Developed using Java, the Eclipse platform can be used to develop rich client applications, integrated development environments and other tools. Eclipse can be used as an IDE for any programming language for which a plug-in is available. Eclipse really is the best IDE for Java development, though it's hard to deny it has been losing ground to IntelliJ recently — perhaps because people don't take the time to look at what Eclipse is capable of. After opening this we can see also the options in the above .click project then a new window will be opened .Then launch the program.right click the one where the program is located and right click it the click the option run as and then click run as java application.It will take some time to run and then as we see in the below image frontend compilation will be started. Now the application will start running succesfully.

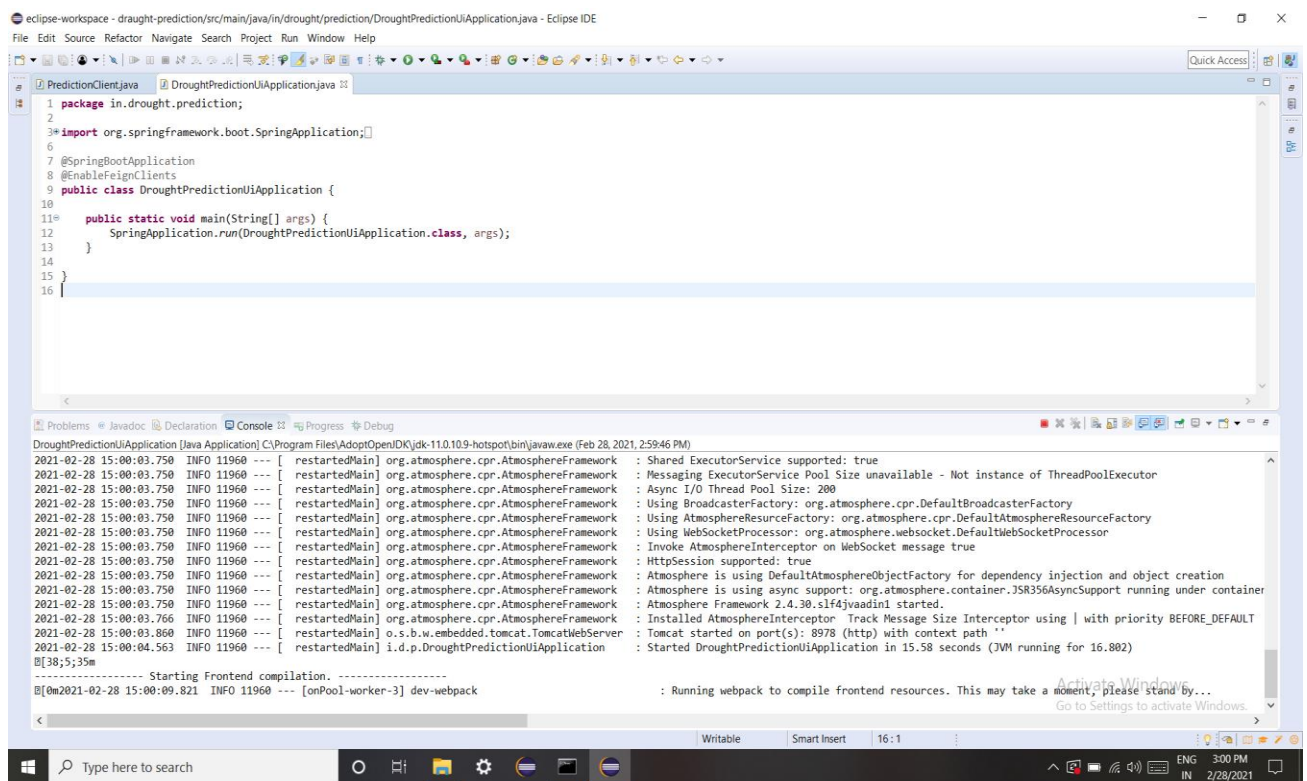


FIG 5.3:Running as java Application

The below screenshot shows that the drought does not occur. we can predict the drought in the app or by using the local port address also we can predict the drought very easily by one click. By giving some values for the pressure, temperature and rainfall and by submitting it we can get the result. The time of the result will depend on the internet connectivity. We need to have a stable internet connectivity. The process in the below screenshot is done in the portal. The same process is to be followed when it is done in the app. By giving the values to the pressure, temperature and the rainfall we can calculate the drought. Developers use the local host to test web applications and programs. Network administrators use the loopback to test network connections. Another use for the localhost is the hosts file, where you can use the loopback to block malicious websites.

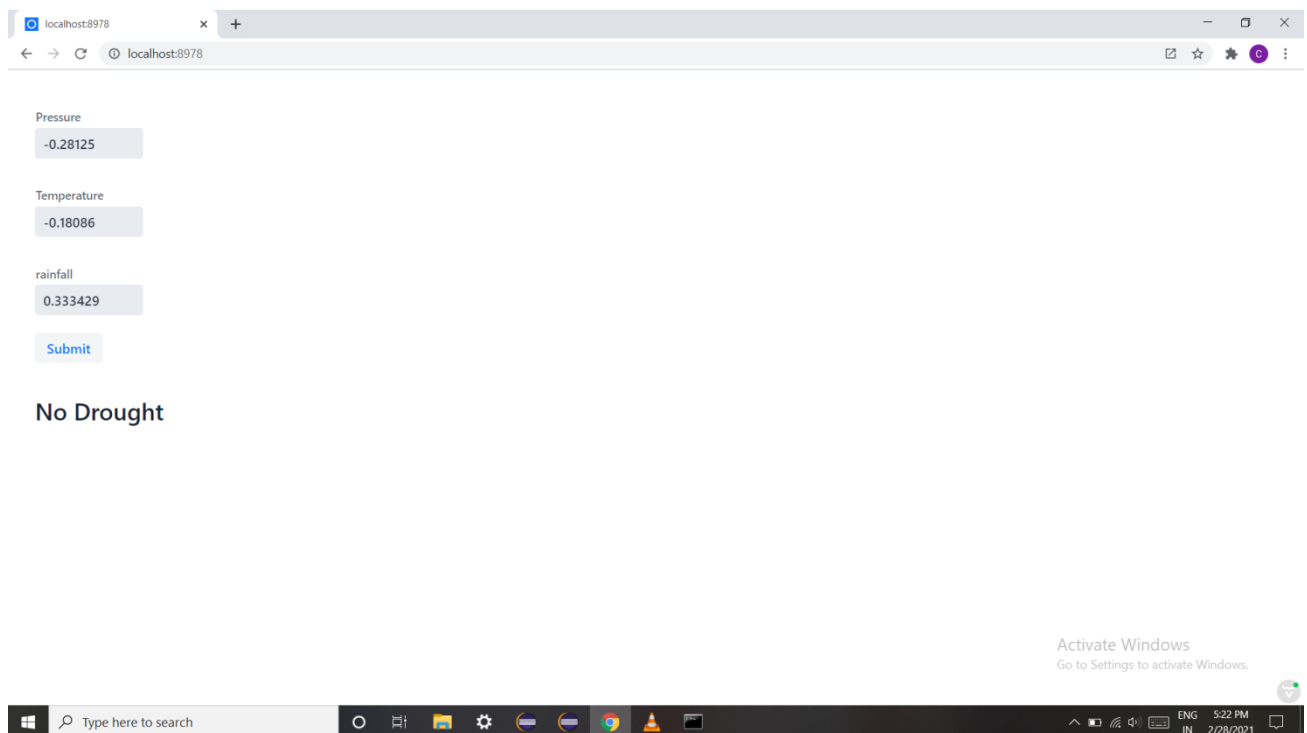


FIG 5.4 Drought does not occur

The below screenshot shows that the drought will occur. we can predict the drought in the app or by using the local port address also we can predict the drought very easily by one click.

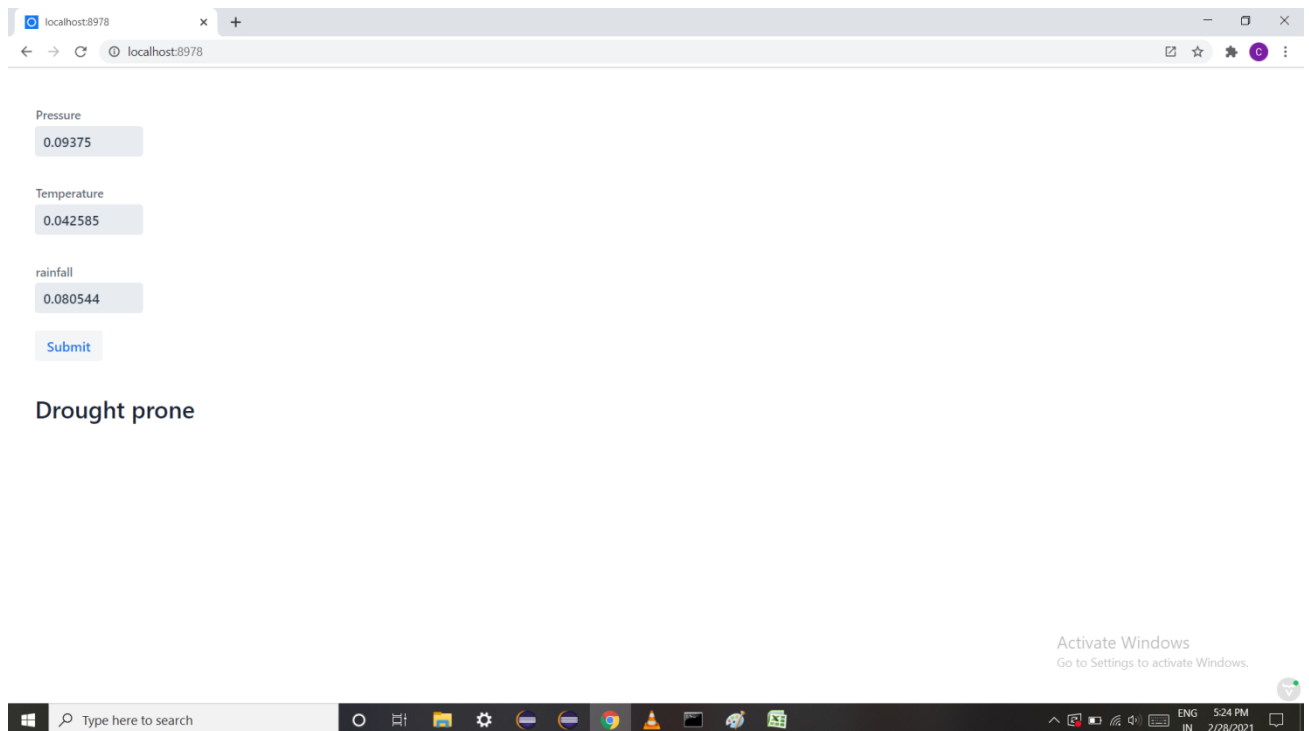


FIG 5.5 Drought occur

CHAPTER 6

CONCLUSION AND FUTURE WORK

This study reviewed recent developments of seasonal drought prediction methods from statistical and dynamical perspectives. The large scale circulation patterns, land surface interactions, and land initial conditions provide important predictability of drought at the seasonal time scale. Meanwhile, a variety of important contributing factors, including local climate, evaporation changes, and human activities may maintain or amplify different types of drought. Seasonal drought prediction mainly focuses on fully utilizing these predictability sources or contributing factors operating at local, continental, or global scales to facilitate drought prediction at seasonal time scales.

For the statistical prediction, these predictability sources and contributing factors are mainly represented as oceanic, atmospheric, or hydrologic predictors and are then fed to a forecast model with the key task to model linear or nonlinear relationships between the predict and and predictors. For the dynamical prediction, GCMs and hydrologic models parameterize physical processes in climatic and hydrologic systems for predicting droughts of different types. Comparison of dynamical and statistical prediction methods reveals the differences in the requirement of computing resources and ability to model complicated interactions. This motivates the integration of both methods to combine drought forecasts from different sources for improving drought prediction.

The final conclusion of the drought prediction of a locality is that this is an easy way to predict the drought by an app. It's less time consuming and an easy app for predicting the drought . This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

To improve drought prediction, efforts are still needed in the understanding of drought mechanisms and predictability in different regions or seasons. Apart from this, several future aspects that may contribute to the improved drought prediction.

REFERENCES

- 1) Short term drought prediction by Anteneh Belanyeh and Jan Malinowski
- 2) Meteorological Drought Prediction Using a Multi-Model Ensemble Approach
- 3) DROUGHT INDICES Andrej Ceglar Biotechnical faculty, University of Ljubljana
- 4) Researchgate.net publication ,238733223 , Methods for predicting drought occurrences
- 5) http://www.dmcsee.org/uploads/file/71_jrc_drought_forecasting_methods_b_kurnik.pdf
- 6) <https://albertawater.com/drought-mitigation/drought-forecasting-and-monitoring>
- 7) <https://drought.unl.edu/Education/DroughtIn-depth/Predicting.aspx>
- 8) <https://www.sciencedirect.com/science/article/pii/S0169809514003858>
- 9) <https://ieeexplore.ieee.org/document/7925314>

APPENDIX

A. Paper AcceptanceMail

B. Journal paper

Enhancing Drought Prediction Methodology of a locality

E. Harichandana, chandanaeluri123@gmail.com, BE-CSE, Sathyabama Institute of science and technology

G. Ramya, ramya.gudipudi22@gmail.com, BE-CSE, Sathyabama Institute of science and technology

ABSTRACT

One of the greatest difficulties confronting the farming framework is the way to defeat the impacts of dry spell. Reducing the effect of dry spell requires a powerful and opportune checking framework. Compelling dry season checking can help build up an early notice framework. Data mining extracts the valuable information from such large data archives in an automated and efficient way. Most of the predictions primarily depends on precipitation rate. Precipitation rate is a major factor in causing drought, but when predicting for a localized area, there are certain other factors which can also have a potential impact on drought. So, this methodology involves basic parameters like humidity, temperature, ground water level, rainfall, wind speed for predicting drought. This is also a less complex method in predicting drought comparing with existing drought predicting methodologies.

Keywords: Drought Predicting, Encryption, Decryption

1. Introduction

A drought is a prolonged shortage in the water supply, whether atmospheric, ground water or surface water in an area. It can substantially affect the biological community of the influenced area and damage to the nearby ecosystem. Continuous drought and water levels below the minimum leading to direct danger for amphibian life.

Drought is an adverse environmental phenomenon that influences almost all aspects of society. It is one of the features of climate and its occurrence is unavoidable.

Of the relative multitude of outrageous climate occasions, dry spell is the most intricate condition influencing numerous individuals. Distinguishing past dry spells and anticipating future dry seasons are imperative to diminish their effect.

However, the arbitrary and whimsical dry spell changeability makes exact dry season figures still a difficult logical test.

This method predicts drought by assessing impact of individual factors on drought, analyzing their relationship in them and draws the conclusion regarding the drought.

1.1 Short term drought prediction by Anteneh Belanyeh and Jan Malinowski

The moderate soil dampness dry season figure is useful to the dynamic of the farming area and the administration of water assets. Part of the ability for moderate dry season anticipating comes from precipitation. Appropriate observing and revision of blustery climate can improve dry spell gauges. Here, we investigated interrelated capacities, which implies that anticipated precipitation can be acquired from the Global Ensemble Forecast System (GEFS) for dry spell gauges in focal India. Climatological mean (CLIM) is utilized as a perception estimate for the assessment of GEFS precipitation figures. The investigation depends on gauges beginning the first and fifteenth day of every month, up to the seventh day. Use relationship and RMSE to assess the measure of GEFS precipitation figures aggregated from 1 to 7 days of advantage. The

volume list dependent on the 2X2 debacle table has been utilized to identify the missed and erroneously assessed every day precipitation

holes in different regions and offices of GEFS. GEFS showed an

improvement in the 0.44 super CLIM compound during the storm season and

0.55 throughout the cold time of year. The RMSE beneath is demonstrated by GEFS rather than CLIM. In the stormy season and winter, the RMSE levels of GEFS and CLIM are 0.82 and 0.4 (total potential is zero), separately. At long last, we utilized there considered GEFS gauge to acquire the Variable Infiltration Capacity (VIC) model, which was utilized to improve the medium-term conjectures for hydrology and agribusiness (soil dampness) in India.

Meteorological Drought Prediction Using a Multi-Model Ensemble Approach

The current NMME SPI anticipating structure is like the CFSv2 SPI determining framework created by Yoon et al. (2012). For each model, first, as indicated by the probability distribution functions (PDF) from the gauge, change the inclination of the month to month precipitation estimate (P), and bias corrected and spatially downscaled (BCSD) to a goal of 0.5 inside the US constant Grid zone network. Thus, the BCSD framework adjusts the climate and general persistent deviations. In particular, for every month and lead time, the PDF for every framework region will be determined dependent on post reposition other than the objective year. From there on, as per the percentile determined from the PDF of the foundation likelihood, the changed rate for the expected year deviation is found in the backwards PDF of the Pexamination. The BCSD strategy is applied to every part and every head of the posterity of P. An updated P conjecture and fused into the coordinated CPC examination (Chen et al., 2008) to frame a P-time arrangement to compute SPI for multi month, 3 months, a half year and a year. The NMME outfit SPI hypothesis is a genuinely clear proportion of six model forecasts. Two operational measures, anomaly correlation coefficient (ACC) and root-mean-square

errors(RMSE) according to perception, are utilized to evaluate prescient abilities. In this investigation, the consolidated CPC environment examination was utilized because of observational environment evaluation.

DROJFI4T fNDIC ES Andre,jCeglar Biotechnical faculty

It dissects the environment of every district and changes the time factor in like manner (dynamic items)

The term factor influences the affectability of the marker to inadequate dampness by deciding the weight given by the current dampness lop-sidedness and the past pointeresteem.

Two arrangements of time coefficients are determined all over the place (one for wet spelling and one for dryspelling)

Climate decency gauge: 2°/c of the PDS1 esteem dips under - 4.0, while 2°/c is above 4.0 => This gives indications of upper and lower limits

2°/cofinordinateddrynessand2°/cof unnecessary wateroccasions

The assessed PDS1 works more reliably than the first PDS1 (Wells et al., 2004)*and gives a precise pointer of dry spell or extraordinary y soil dampness.

2.EXISTING SYSTEM

The existing system for drought forecasting uses many techniques like data mining, machine learning and deeplearning.

The technique of data mining is introduced to find associations between drought and several oceanic and climatic indices and thus predicting thedrought.

Machine learning techniques hav'e likewise as of late been utilized to anticipate dry season, utilizing dry spell pointers, for example, the Standardized Precipitation Index (SPI) as information passage.elearning methods, like the artificial neural networks(ANN), supJxirtvector regression (SVR) and incor}xirated ANN wavelet, use wavelet examination to handle input information and utilize two execution markers to think about anticipating aftereffects of these there are three different ways.

Drought can be anticipated utilizing machine readings dependent on distant information detectinginformation.Alternatively, the Deep Belief networkis comprised of two Boltzmann limited machines, which are utilized to utilize the rest of the Standardized Stream flow Index (SSI) as an indicator of long- haul dry spell. This technique is reasonable for anticipating dry season markers on different time scales.

2.1 DISADVANTAGES OF EXISTING SYSTEM

Individual assessment of each and every causative parameter is not involved in the existingmethods.

There is a need for a tool to predict the water scarcity in an area, sometimes as small as size of a range of 1000ha.Most of the predictions primarily depends on precipitation rate. Precipitation rate is a major factor in causing drought but when predicting for a localized area there are certain other factors which can also have a potential impact on drought

3.PROPOSED SYSTEM

The proposed plan assesses the causes of drought and predicts the risk of drought.If the occurrence of the drought is detected, then the severity of the drought is determined.

If a drought occurs, then the size will clear.Machine learning algorithm is used for analytic purpose.

By This is done by analyzing the historical value of the causal parameter and predicting the future value.

of drought. Thus, Machine learning algorithms are used for analysis purposes.

Therefore, the predicted value is tested todetermine the probability

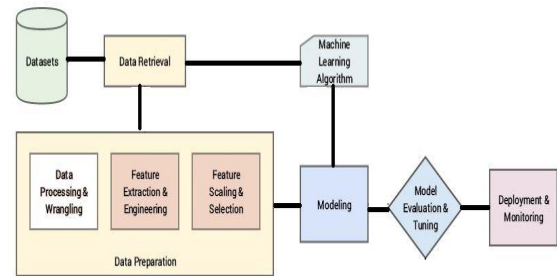


Fig 1.1 System Architecture

3.1 SYSTEM REQUIREMENTS:These are the software and hardware requirements that are used for system installation.

3.1.1 HARDWARE REQUIREMENTS:

System - Pentium-IV
Speed - 2.4GHZ
Hard disk - 40GB
Monitor - 15VGA color
RAM - 512MB

3.1.2 SOFTWARE REQUIREMENTS:

Operating System - Windows XP
Coding language - Java
IDE - Net beans
Database -MySQL

3.2MODULES

3.2.1.Analysis Module
3.2.2.Drought Prediction Module
3.2.3.Calculate Severity index
3.2.4.View Information Module

3.2.1 Analysis Module

In this Module, analysing the historical values of the causative parameters and predicting the upcoming values. Machine learning algorithm is used for analytic purpose.

3.2.2 Drought Prediction Module

This method predicts drought by assessing impact of individual factors on drought, analyzethe relationship in them and draws the conclusion regarding the drought. Most of the predictions primarily depends on precipitation rate, and basic parameters like humidity, temperature, ground water level, rainfall, wind speed for predicting drought. This is also a less complex method in predicting drought comparing with existing drought predicting methodologies.

3.2.3 Calculate Severity Index

In this Module, Calculate the severity index based on drought prediction parameters like humidity, temperature, ground water level, rainfall, wind speed for predicting drought.

3.2.4 View information Module

In this Module, user can view the information about how many percentage drought will be occurred.

3.3 USECASE

Use case diagrams overview the usage requirement for system. They are useful for presentations to management and/or project stakeholders, but for actual development you will find that use cases provide significantly more value because they describe “the meant” of the actual requirements. A use case describes a sequence of action that provides something of measurable value to an action and is drawn as a horizontal ellipse.

This study checked on the most recent improvements in yearly dry spell conjecture strategies from a measurable and energy point of view. Broad pivot designs, earth associations and introductory earth conditions give huge estimates of dry spell in the occasional scale. Simultaneously, an assortment of contributing variables, including neighbourhood environment, changes in vanishing and human movement, can support or fuel various sorts of dry spell. Foreseeing the yearly dry spell predominantly centres around utilizing these expected sources or impacting factors on the neighbourhood, mainland or worldwide scale to advance dry spell gauges on occasional scales. Numerical estimates, these speculative sources and powerful factors are introduced fundamentally as marine, rocket or prescient hydrological articles, and afterward consolidated into the prescient model. Its primary capacity is to impersonate the immediate or non-direct connection among theoretical and prescient gadgets. With solid hypothesis, GCM and hydrological boundary models are an interaction of environment and hydrological framework to foresee different kinds of dry spell. Examinations of ground-breaking expectations and measurements uncover the fundamental contrasts in PC administrations and the capacity to show complex associations. This animates the combination of two strategies for consolidating dry spell gauges from various sources to improve dry season figures.

4.RESULTS

This method predicts drought by assessing impact of individual factors on drought,analyzethe relationship in them and draws the conclusion regarding the drought.Analysing and Predicting the drought based on humidity, temperature, ground water level, rainfall, wind speed.

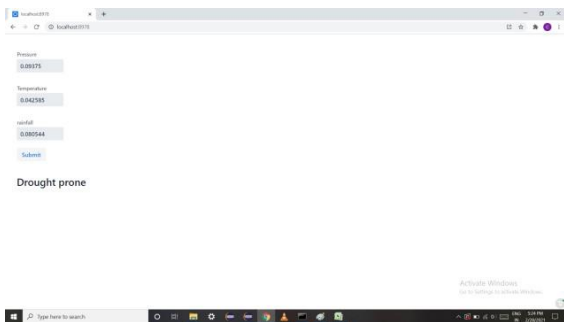


Fig 1.2 Drought Prone

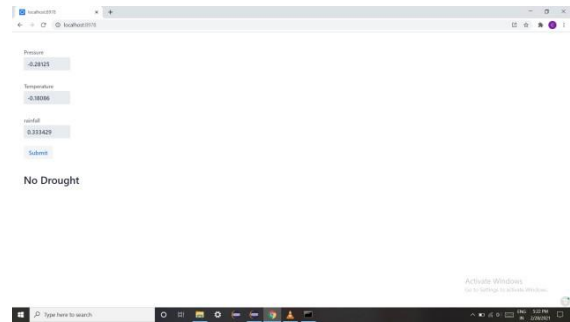


Fig 1.3 Drought does not occur

5.FUTURE SCOPE

To improve drought prediction, efforts are still needed in the understanding of drought mechanisms and predictability in different regions or seasons. Apart from this, several future aspects that may contribute to the improved drought prediction.

6.OUTPUTGRAPH

Here boolean values are used for output.if the result is 0 then the drought does not occur.if the result is 1 then the drought occurs.

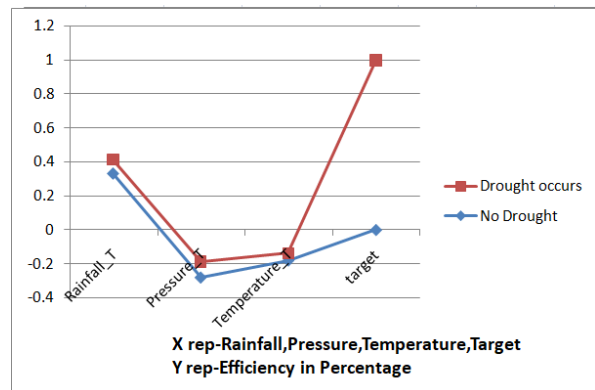


Fig 1.4 Output Graph

7.REFERENCES

- [1] Short term drought prediction byAntenehBelanyeh and Jan Malinowski
- [2] Meteorological Drought Prediction Using a Multi-Model Ensemble Approach
- [3] DROUGHT INDICES Andrej Ceglar Biotechnical faculty, University of Ljubljana
- [4] Researchgate.net publication ,238733223 , Methods for predicting drought occurrences
- [5] http://www.dmcsee.org/uploads/file/71_jrc_drought_forecasting_methods_b_kurnik.pdf
- [6] <https://albertawater.com/drought-mitigation/drought-forecasting-and-monitoring>
- [7] <https://drought.unl.edu/Education/DroughtIn-depth/Predicting.aspx>
- [8] <https://www.sciencedirect.com/science/article/pii/S0169809514003858>
- [9] <https://ieeexplore.ieee.org/document/7925314>

D. Signed Bonafide:

E. Source Code:

PREPROCESSING

```
import pandas as pd
from sklearn.preprocessing
import MinMaxScaler
from sklearn.model_selection
import train_test_split
from sklearn.preprocessing
import
LabelEncoder,OneHotEncoder
```

```
class scaler(object):
```

```
def __init__(self):
    pass
```

```
def scaling(self,x):
    sc=MinMaxScaler(feature_range
    =(0,1))
    scaled_data=sc.fit_transform(x)
    return(scaled_data)
```

```
    return(y)
```

```
def __del__(self):
    pass
```

MAIN FORECAST

```
#!/usr/bin/env python
# coding: utf-8
```

```
# In[81]:
```

```
from bs4 import BeautifulSoup
import urllib.request
from time import sleep
from datetime import datetime
import pandas as pd
import requests
import re
from datetime import date
import numpy as np
import matplotlib.pyplot as plt
```

```
def __del__(self):
    pass
```

```
class splitter(object):
```

```
def __init__(self):
    pass
```

```
def
decomposition(self,x,y,test_size=
0.25,random_state=0):
```

```
x_train,x_test,y_train,y_test=train
_test_split(x,y,test_size=test_siz
e,random_state=random_state)

return(x_train,x_test,y_train,y_te
st)
```

```
import seaborn as sb
```

```
# In[82]:
```

```
def getweather():
    weather = []
    year = date.today().year
    url =
"http://www.hko.gov.hk/cis/dailyE
xtract/dailyExtract_" + str(year) +
"08.xml"
```

```
    page = requests.get(url)
```

```
    soup =
BeautifulSoup(page.content,
'html')
```

```
    body = soup.find("body").text
    body = body.split(",")
    # weather.append(temp)
    day = body[-27][4:-1]
    if (day[0] == "0"):
```

```
def __del__(self):
    pass
```

```
class Encoder(object):
```

```
def __init__(self):
    pass
```

```
def labelEncoding(self,x):
    le=LabelEncoder()
    x=le.fit_transform(x)
    return(x)
```

```
def one_hot_encoder(self,y,cf):
    ohe=OneHotEncoder(categorical
_features=[cf])
    ohe=OneHotEncoder(categorical
_features=cf)
```

```
y=ohe.fit_transform(y).toarray()
    day_2 = day[-1]
    else:
```

```
    day_2 = day
    weather.append(day_2)
    month = body[0][-1]
    weather.append(month)
    weather.append(year)
```

```
    temp = body[-24][1:-1]
    weather.append(temp)
```

```
    high = body[-25][1:-1]
```

```
    low = body[-23][1:-1]
    weather.append(high)
    weather.append(low)
```

```
    Humidity = body[-21][1:-1]
    weather.append(Humidity)
    dew = body[-22][1:-1]
    weather.append(dew)
    pressure = body[-26][1:-1]
    weather.append(pressure)
    if (float(temp) >= 30):
        heat = "YES"
```

```

else:
    heat = "NO"
    wet = "NO"
weather.append(wet)

print("Day : ", day_2)
print("Month : ", body[0][-1])
print("Year : ", year)
print("Pressure : ", body[-26][1:-1])
print("Max Temp : ", body[-25][1:-1])
print("Mean Temp : ", body[-24][1:-1])
print("Min Temp : ", body[-23][1:-1])
print("Dew Point : ", body[-22][1:-1])
print("Humidity : ", body[-21][1:-1])
print("Heat : ", heat)
print("Wet : ", wet)
    # print("Temperature : ",temp)

    return weather

# In[83]:

getweather()

# In[84]:

import csv
import openpyxl
import sys

def main():
    print("Collecting Weather Data
    in a intervals of 30 mins")
    idx = 0
    weatherdata =
    {'Day':[], 'Month':[], 'Year':[], 'mean
    _temp':[], 'max_temp':[], 'min_tem
    p':[], 'meanhum':[], 'meandew':[], 'pr
    essure':[], 'heat':[], 'wet':[]}
    sc=scaler()
    en=Encoder()

```

```

weather.append(heat)
    if (float(Humidity) >= 80):
        filename = 'data.xlsx'
wb =
openpyxl.load_workbook(filename
e=filename)
    sheet = wb['Sheet1']
    #new_row =
['Day','Month','Year','mean_temp'
,'max_temp','min_temp','meanhu
m','meandew','pressure','heat','w
et']
    #sheet.append(new_row)
    while idx< 1:
print('collecting weather data
'+str(idx))
tmp = getweather()
weatherdata['Day'].append(tmp[0
])
weatherdata['Month'].append(tm
p[1])
weatherdata['Year'].append(tmp[
2])
weatherdata['mean_temp'].appe
nd(tmp[3])
weatherdata['max_temp'].append
(tmp[4])
weatherdata['min_temp'].append
(tmp[5])
weatherdata['meanhum'].append
(tmp[6])
weatherdata['meandew'].append(
tmp[7])
weatherdata['pressure'].append(
tmp[8])
weatherdata['heat'].append(tmp[
9])
weatherdata['wet'].append(tmp[1
0])

try :
    new_row =
[int(tmp[0]),int(tmp[1]),int(tmp[2]),
float(tmp[3]),float(tmp[4]),float(tm
p[5]),float(tmp[6]),float(tmp[7]),flo
at(tmp[8]),tmp[9],tmp[10]]
    sheet.append(new_row)
    except ValueError as e:

# In[88]:

```

```

        wet = "YES"
    else:
        print (e)
idx+=1

wb.save(filename)

    print(weatherdata)

# In[85]:

main()

# In[86]:

import pandas as pd
data_xls =
pd.read_excel('data.xlsx',
'Sheet1', index_col=None)
data_xls.to_csv('data.csv',
encoding='utf-8', index=False)

# In[87]:

#importing classes to handle
data,to split data into training
and testing sets,tovisualising of
tree and to cheak accuracy of
our model
from preprocessing import
splitter
from preprocessing import
Encoder
from preprocessing import scaler
from metrics import matrix
from treeVis import vis
sp=splitter()
mt=matrix()
vs=vis()

#importing dataset
dataset=pd.read_csv("data.csv")

```

```
dataset.tail()
```

```
# In[89]:
```

```
from sklearn.preprocessing
import LabelEncoder
lab=LabelEncoder()
dataset.iloc[:,0]=lab.fit_transform
(dataset.iloc[:,0])
dataset.iloc[:,1]=lab.fit_transform
(dataset.iloc[:,1])
dataset.iloc[:,2]=lab.fit_transform
(dataset.iloc[:,2])
dataset.iloc[:,3]=lab.fit_transform
(dataset.iloc[:,3])
dataset.iloc[:,4]=lab.fit_transform
(dataset.iloc[:,4])
dataset.iloc[:,5]=lab.fit_transform
(dataset.iloc[:,5])
dataset.iloc[:,6]=lab.fit_transform
(dataset.iloc[:,6])
dataset.iloc[:,7]=lab.fit_transform
(dataset.iloc[:,7])
dataset.iloc[:,8]=lab.fit_transform
(dataset.iloc[:,8])
dataset.iloc[:,9]=lab.fit_transform
(dataset.iloc[:,9])
dataset.iloc[:,10]=lab.fit_transfor
m(dataset.iloc[:,10])
```

```
dataset.tail()
```

```
# In[90]:
```

```
#deviding data into dependant
and independant sets
x = dataset[['Day','Month','Year']]
y = dataset['heat']
z = dataset['wet']
```

```
# In[91]:
```

```
#visualising each attribute of
dataset using histogram
dataset.hist(figsize = (30, 30))
```

```
plt.savefig("dataset.png")
plt.show()
```

```
# In[92]:
```

```
#deviding data into training and
testing sets
from sklearn.model_selection
import train_test_split
x_train,x_test,y_train,y_test=train
_test_split(x,y,test_size=0.20,ran
dom_state=42)
x_train,x_test,z_train,z_test=train
_test_split(x,z,test_size=0.20,ran
dom_state=42)
```

```
# In[93]:
```

```
#importing decision tree model
and fitting training data to it
from sklearn.tree import
DecisionTreeClassifier
classifier=DecisionTreeClassifier
(criterion='entropy',random_state
=0)
classifier.fit(x,y)
```

```
classifier_z=DecisionTreeClassifi
er(criterion='entropy',random_sta
te=0)
classifier_z.fit(x,z)
```

```
# In[94]:
```

```
#predicting values for x_test and
compairing result with y_test
print("Heat:-")
y_pred=classifier.predict(x_test)
print("predicted values:-")
print(y_pred.astype(int))
y_test_arr=np.array(y_test)
print("original values:-")
print(y_test_arr)
```

```
#predicting values for x_test and
```

```
compairing result with y_test
print("Wet:-")
z_pred=classifier_z.predict(x_tes
t)
print("predicted values:-")
print(z_pred.astype(int))
z_test_arr=np.array(z_test)
print("original values:-")
print(z_test_arr)
```

```
# In[95]:
```

```
#cheaking accuracy of our model
accuracy=mt.accuracy(y_pred,y_
test)
print("{}%".format(accuracy*100))
```

```
accuracy=mt.accuracy(z_pred,z_
test)
print("{}%".format(accuracy*100))
```

```
# In[96]:
```

```
# In[99]:
```

```
from datetime import date
now = date.today()
filepath = "prediction/forecast-
"+str(now)+".xlsx"
def prediction():
wb = openpyxl.Workbook()
wb.save(filepath)
print("Predict Weather Data ")
idx = 0
weatherdata =
{'Day':[],'Month':[],'Year':[],'heat':[]
,'wet':[]}
filename = filepath
wb =
openpyxl.load_workbook(filenam
e=filename)
sheet = wb['Sheet']
new_row =
['Day','Month','Year','Heat','Wet']
```

```

sheet.append(new_row)
    while idx< 31:
print('Predict weather data ' +
str(idx))
    year = date.today().year
yr = year - 1999
    if (date.today().day == 31):
        year += 1
yr += 1
y_pred = classifier.predict([[idx,
0, yr]])
    print(y_pred.astype(int))
z_pred =
classifier_z.predict([[idx, 0, yr]])
    print(z_pred.astype(int))
tmp_d = idx + 1
tmp_m = 8
tmp_y = year
tmp_h = y_pred.astype(int)
tmp_w = z_pred.astype(int)
    if (int(tmp_h) == 0):
tmp_h_char = "NO"
    else:
tmp_h_char = "YES"
    if (int(tmp_w) == 0):
tmp_w_char = "NO"
    else:
tmp_w_char = "YES"
new_row = [int(tmp_d),
int(tmp_m), int(tmp_y),
tmp_h_char, tmp_w_char]
sheet.append(new_row)

```

```

idx += 1
wb.save(filename)

    print(weatherdata)

# In[100]:

prediction()

# In[101]:

data_xls =
pd.read_excel(filepath, 'Sheet',
index_col=None)
data_xls.to_csv("prediction/forec
ast-"+str(now)+".csv",
encoding='utf-8', index=False)

# In[102]:

#importing dataset
data_pred=pd.read_csv("predicti
on/forecast-"+str(now)+".csv")
data_pred

```

```

# In[103]:

import seaborn as sns
sns_plot =
sns.countplot(x='Heat',data=data
_pred)
fig = sns_plot.get_figure()
fig.savefig("prediction/heat_sum
mary_prediction-
"+str(now)+".png")

# In[104]:

sns_plot =
sns.countplot(x='Wet',data=data
_pred)
fig = sns_plot.get_figure()
fig.savefig("prediction/wet_summ
ary_prediction-
"+str(now)+".png")

# In[105]:

```