IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 5, Ver. I (Sep.- Oct. 2017), PP 01-10
www.iosrjournals.org

The Open Environment System: Rainfall Prediction Using Naïve Bayesian Algorithm

*Anik Pait[1], Venkatesan M[2]
[1]*B.Tech Department of Computer Science and Engineering, VIT University, Vellore, Tamil Nadu, India*
[2]*Associate Professor Department of Computer Science and Engineering, VIT University, Vellore, Tamil Nadu, India*
Corresponding Author: Anik Pait

Abstract: *There has been a tremendous growth in data over the years. The question often lies in how the data is being used on a daily basis to predict the uncertainties. If we look at it from a climatic change perspective, though we have a lot of sources of data to be informed about the weather prediction and what if, a cost effective, simple and portable method is also available for the same. Hence the idea of "Open Environment System" is being proposed which is a prototype that will predict the possibility of the rainfall on a given day as well as can predict the minimum & maximum rainfall possible in that region on the particular month, based on the factors that cause rain along with the availability of the historical data.*

Date of Submission: 29-07-2017 Date of acceptance: 08-09-2017

I. Introduction

Rain always carries uncertainties with itself which are often unpredictable. But the significance of such uncertainties serves answers to many questions. As we know, rain is an important asset for groundwater, irrigation practices by farmers for the crops, for prediction of floods and landslides. But people often struggle to get the right prediction of the rainfall and hence it can raise problems in decision making where rain plays a critical role. If the right amount of rainfall is not predicted then one can imagine that how such ambiguous predictions can be hazardous for the people. Though there are many systems which exist to predict the rainfall and these systems might work with different algorithms but such systems are either very expensive or the algorithms are very complex. Hence, the introduction of the Open Environment system brings in a cost effective approach which can predict the amount of rainfall possible for a particular day which makes it useful for everyone not only in terms of cost but also for re-usability. At present, due to the huge boom in the technological sector, there are many prototypes available in the market that can easily give the temperature and humidity reading and even rainfall prediction but they have issues with it. There are also plenty of devices including smart phones which give data directly from the meteorological department but internet connection or other forms of communication services must be in place. This could pose a problem for those living in remote areas or might have a tough time working with sophisticated technologies or might not even have the resources to buy or maintain them, these could be farmers, tribal people, fisherman etc. A cost effective method that has portability, simple usage, and easy maintenance can reduce the load from people in their daily lives and make their lives simpler.

It will predict the possibility of the rainfall on a particular day as well as it will predict the minimum and the maximum rainfall possible in that region based on the historical data. The Arduino Board will have the required sensors to measure the temperature and humidity as they are the main factors which will decide the forecast of the rain. Only two of the important factors which are humidity and temperature have been taken into account for the study. The temperature and humidity sensors will gather the required data of the particular day from the surroundings which will be the put for the microcontroller and based on the input; the system will decide to use the data for the prediction of rain for that particular day itself. It will display the result of prediction instantaneously as a "yes" or a "no" by the Led indicator and the LCD screen display based on the decision algorithm in Java programming.

A database of the historical data about rainfall condition has to be collected for the target region has to be maintained for the records to get an idea whether the assumptions made by us reconcile with the records or not. The C/Java program shapes up the Naïve-Bayesian algorithm so that it can conclude to a decision. The Open

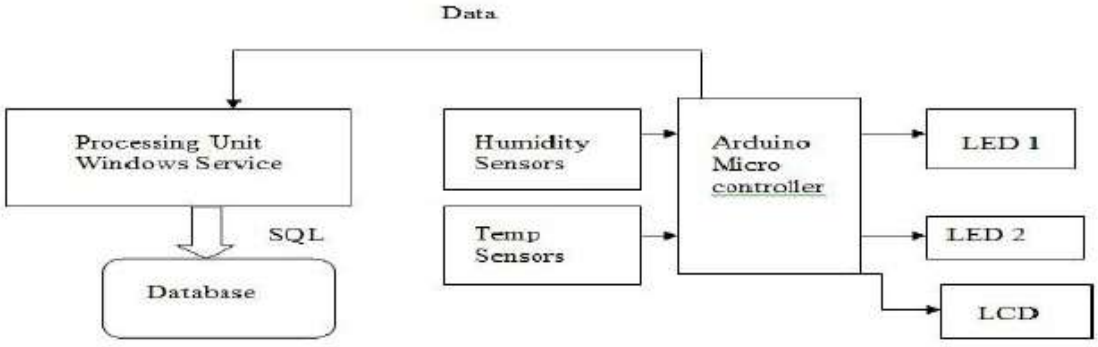
Environment System serves the apt for an additional purpose as well is that if needed, the system can be modified into various other systems like Pollution Detector which decides whether the city can be said to be polluted or not, based on the pollutant levels estimated by the system. The modified system can be produced by doing some changes in the code as well as to the sensors. Hence, the system is not at all stringent. That's why it has been termed as Open Environment system since it can be morphed into any other applications as per need.
Open Environment system can be morphed into many other systems as mentioned n the above statement, we have come across an example of Pollution detection system where based on the data collected from sensors and a decision algorithm, a binary result can be displayed. Similarly, some other examples include Noise processing system where a judgment can be made about the noise pollution as is it crossing the threshold level or not, extreme weather limits detection which needs to consider numerous data/factors to make a judgment on the effects of global warming of that particular area. But the historical data collection can a daunting task. Also, the system will give the range of intensities; the rainfall is possible between min-max, but it won't give the measurement of that particular day (intensity of rain)

1.1 Aim /Objective of the proposed work

- The broad motive of the system is to create an Open Environment System that records and stores the temperature and humidity values and applies the Naïve-Bayesian Classifier algorithm to predict the probability of rain.
- Apart from predicting the rainfall, it will also determine the maximum and minimum intensity of the rainfall for the particular month if it rains on that particular month.
- The max and min intensity will be based on the historical data of rainfall which will be stored in the database.
- The issue for the people is that they tend to plan an excursion with an expectation of sunny mornings & day, or the farmers are highly dependent on rain for their crop cultivation and irrigation practices so that others can get food, rains are the central cause behind floods or landslides which leads to severe damage and loss of lives and property. So the proposed system aims to abate the impact of the harmful effects of rainfall and also to contribute to betterment of the people. If people are already aware of the rainfall they can take the necessary steps required for irrigation practices or protection from harmful diseases or precautions required for floods or landslides because it will make the user familiarize with the intensity of the rainfall one can expect in that particular month which will lie somewhere between the min and max intensity range and will be mathematically calculated.
- The database will store the necessary training data based on the assumptions and another database will be maintained which will store the precipitation levels of that particular city for the whole year which will help for the intensity calculation of the rainfall and hence one will be aware of the rainfall prediction along with its intensity so as to take the necessary preventive steps.

II. Framework Of The Proposed System

Figure 1: Concept Sketch of the proposed System



The Open Environment System consists of an Arduino Uno which will work as a microcontroller which will be connected with two sensors which will record temperature and humidity which are:

- Humidity Sensor Module SY-HS-220 [1]
- LM158/LM258/LM35/LM2904 Low Power Dual Operational Amplifiers. It operates in the range -55 to 120 degree Celsius. [2]

These two sensors will work as an input for the system along with the power supply which is usually the UART cable connected to COM9 in the laptop running on Windows 7/8 OS. As soon as the user makes the connection, the values are recorded by the sensors gets displayed by the 2X14 LCD screen on the dot board. The code file needs to be opened and right clicked to run. A login screen on the web-page pops up asking the user

for the username and password, the user needs to fill it up with relevant details. Once the Username and password are verified by the database, the user will be redirected to the main primary database of the rain forecasting system that will have humidity and temperature values. The verification of the user is done only once and from next time authentication might not be required. The current focus lies on embedding the code into a chip to fit into the system so it runs automatically rather than having a separate PC/Laptop to run it. The values obtained from the sensors are entered as an input to the Java program of Naïve-Bayesian Classifier.
The data collected is processed and the Naïve-Bayesian Classifier Operation is applied to predict the probability of rainfall. The equation for which is given below where:

$\mathbf{X} = (\mathbf{x_1}, \ldots, \mathbf{x_n})$
representing some *n* features (independent variables), it assigns instance probabilities

$\mathbf{P}(\mathbf{C_K}|\mathbf{x_1}, \ldots, \mathbf{x_n})$
for each of K possible outcomes or classes.

$$P(C_K|X) = \frac{P(C_K)P(X|C_K)}{P(X)}$$

III. Literature Review

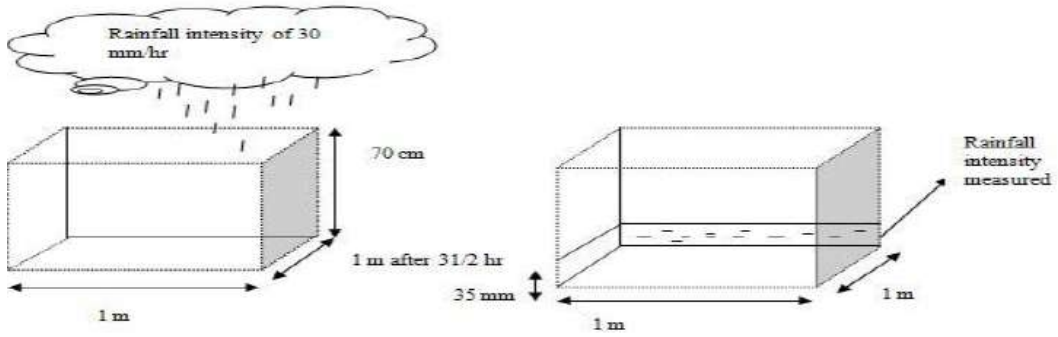
The human forecasters had been responsible for evaluation and generation of entire weather forecast based solely on available observations in the past. The ancient method of weather forecasting was dependent on human experience and ability to observe patterns [3]. For example, if the present day has a sunset, the following day was predicted to have a fair weather [3]. If it was visibly cloudy on a particular day, the next day is predicted for forecasting of rainfall. In the 20th century, the meteorology had made a significant improvement in weather predictions as it has allowed understanding of atmospheric process and invention of the barometer has made a significant impact. The idea of numeral calculations for weather prediction was first brought into the limelight by Lewis Richardson in 1922 [3]. The prediction made by him considered the change in atmospheric pressure for that particular region and time though it failed due to hindrance caused by other factors as the value he got was too unrealistic [6]. However, as of today, the manual human input is majorly restricted to selecting a model based on certain essential parameters, such as model biases and performance. The human is now needed to observe and infer from the model data into weather forecasts that are meaningful and understandable to the end user. Humans may utilize the knowledge of local effects and situation which may be too minuscule to be efficiently resolved by the model to add information to the forecast. While improving the accuracy of prediction models implies that human beings shall no longer be required in the forecast process at a particular point in the future, there is currently still an essential requirement for manual human intervention.
In 2013, an attempt was made for weather prediction system by the scientists where an indigenous rain forecasting system was developed at the Indian Institute of Tropical Management (IITM) [4] to prove its effectiveness in helping people prepare for the unpredictable monsoon. The system, developed by IITM to forecast rain could predict rainfall for 15-20 days after every five days, as was claimed by the scientists at the Institute [4]. They also claimed that it was currently at an experimental stage. It was stated by the Senior IITM scientist Mr. A.K.Sahai that the system which was validated by the Indian Meteorological Department, it had a significant potential to become an indigenous model for rain forecast in the country [4]. The system was expected to be operationalised in the next six months as claimed by the scientists of the Institute following a nod from IMD. Scientists believed that the system of forecasting weather over ten days in advance, had referred to as Extended Range Prediction, could play a proficient role in agricultural planning, dam management, city planning and management that depends on the uncertainty of rainfall of similar resources.

IITM, in 2013 itself, launched an initiative called SAFAR (System for Air Quality and Weather Forecasting) [5] for getting valuable information for disaster management, it helped in weather prediction and monitored air quality in polluted cities. It monitored ultraviolet radiations to keep track of pollution levels [5]. Using the pollution level data, it helped the people in the city to avoid the most polluted regions or to take precautionary measures, especially for the breathing disorder patients. The data was measured by taking different parameters like ozone, mercury, and nitrogen [5]. This was a similar initiative which was launched for Rain forecasting system in India. Apart from these, significant works are going on correspondingly regarding the rain forecasting which might still take the time to release.

3.1 Traditional way of measuring intensity

The role of intensity & how it is measured:

Figure 2: Intensity measurement



In the above figure 2, the most common method (in homes, etc.) has been shown where one can measure the intensity of rainfall by just keeping a box of a given size. During the shower, the container collects the water [7], and after some time, one can take the appropriate measurements for the height of water which can be concluded as the current intensity of the rainfall today. Suppose the shower in the box is measured as 30mm for an hour, then measurement of intensity will be 30mm/hour.[7]

There are numerous new ways to measure the rainfall intensity if it rains. Some of the ways are either by calculating the intensity from the rational method or by creating the Intensity Duration Frequency curve, or there are various formulas made by the hydrologists to measure the same.

3.1.1 Time of Concentration

The time of concentration is defined as the time taken to travel from the most hydrological remote point to the point where the water is collected where travelling is done by the water particle is called time of concentration [8]. This can also be termed as max peak off rate from a small area due to rain that lasts long enough so that there is a concentration of runoff from all parts of the watershed and hence the time required for the collection is called time of concentration(Tc).

By using the equations to determine the rainfall intensity for computing peak discharge results in an iterative process because the rainfall intensity itself is a function of the time of concentration. So if there is an inverse relationship; between the time of concentration and the rainfall intensity which makes it one of the key factors to use for rainfall intensity and hence we can determine the rainfall intensity by using the time of concentration. A preferable range used for the time of concentration is from 5 minutes to 60 minutes which acts as the range of time of concentration. Any value in between can be used, but the first and the last value can be the extremes for the range of rainfall intensities.

IV. Experiment: Using The System

The system was 'tested' in Vellore where the sensors recorded the values as:
Temperature=32 degree Celsius
Humidity=36%
'Training Data' has following fields/attributes on which prediction is applied:-
1) Outlook : (sunny/overcast/rainy)
2) Temperature : (hot/cool/mild)
3) Humidity : (normal/high)
5) Class/Rain : (yes/no)
Based on a series of 'if-else' conditions we have assumed if the temperature is >30, it will be a 'hot' temp category. (Based on the past observations and the common reaction by the people that how they used to categorize the temperatures & humidity, the assumptions have been made). Based on the if-else conditions considering both temp and humidity, Outlook has been categorized/assigned.

- If(temp=='Hot' && Humidity=='High')

then Outlook='Sunny'.
(Java program)
Similarly, we will have different 'if conditions' for different scenarios of temperature and humidity. Separate Coding files have made to assign these conditions.
Since temp=32 implies:
Temp: Hot
Since Humidity=36% implies:

Humidity: Normal
If (Temp==”hot” && Humidity==”Normal”)
Then Outlook=”sunny”
Coding file (if-else) Purpose: To assign Temp as ‘Hot/Cool/Mild’
To assign Humidity as ‘High/Normal’
(Based on the Sensor Readings)
To assign Outlook parameter based on the first two conditions of temp and humidity.
Coding file (Arduino) purpose: To enter the final probability values of rainfall or no rainfall to the portable system to display the final result in the form of LED light.

4.1 Algorithms used for Decision Step
Algorithms used for determining the rainfall and its intensity are the Naïve Bayesian algorithm and the rational method is used for intensity.

 Naïve Bayesian Algorithm (Manual Calculations)
The Naïve Bayesian algorithm is used as a prediction algorithm based on the training data after which the testing data is tested, and a decision is reached finally [9].
 We have two classes “yes” => It will rain and “No”=> It will not rain. Always we need to identify thenumber of classes first.
Past Cases:
Let C1: it has rained (yes)
C2: it has not rain (No)
P(C1)=3/5=0.6
P(C2)=2/5=0.4 (P=> Probability)
 Enter the Test Case as an input to the Naïve Bayesian Classifier which will take one input for each of the parameters/attributes to which the decision attribute is dependent on [9].

```
mysql> use rain;
Database changed
mysql> select * from Rain;
+----------+------+----------+-------+
| outlook  | temp | humidity | class |
+----------+------+----------+-------+
| sunny    | hot  | high     | no    |
| overcast | cool | normal   | yes   |
| rainy    | cool | normal   | yes   |
| rainy    | mild | high     | no    |
| sunny    | mild | normal   | yes   |
+----------+------+----------+-------+
5 rows in set (2.23 sec)
```

The database has all the parameters –Outlook, Temp & Humidity and the classes(Rain-yes or no) are based on assumptions (specifically on outlook, for example, if the outlook is sunny=> it will not rain (based on past references & common logic/Observed trait over the years)) and so on.
 Similarly, the previous step is done for each attribute record with the training data and the decision of the class attached to that particular record.

Let N=> Count the presence in database
‘/’=> Provided
total=> Total number of Yes in the case of P (it will rain) and the total number of No in the case of P (no rain)
N(temp=”Hot”/C1) =0
N(temp=”Hot”/C2) =1
and similarly, calculating for humidity & Outlook:
N(humidity=”Normal”/C1)=1 & N(humidity=”Normal”/C2)=2
N(Outlook=”Sunny”/C1)=1
N(Outlook=”Sunny”/C2)=1
 Apply the multiplication for each test case, the combination with its decision and get the result as the probability of whether the decision will be yes or no.

P(it will Rain)= (N(Outlook=”Sunny”/C1)/total* N(temp=”Hot”/C1)/total* N(Humidity=”normal”/C1)/total *P(C1))/total = 0 Since N(temp=”Hot”/C1=0)
P(it will not Rain)=(N(Outlook=”Sunny”/C2)/total* N(temp=”Hot”/C2)/total*
N(Humidity=”normal”/C2)/total*P(C2))/total =1/2*1/2*2/2*0.4

 The Probability of 'yes' implies it's going to rain and 'no' implies it's not going to rain. The decision is compared and the result is crystal clear.

Since P (it will Rain)=0 which is not a valid quantity in this case=> Applying Laplace estimator,
Total=Total +2 (Classes)=3+2=5
Add 1to the count=>
P(it will Rain)= (P(Outlook="Sunny"/C1)* P(temp="Hot"/C1)* P(Humidity="normal"/C1) *P(C1))
=(0+1)/5*(1+1)/5*(1+1)/5*0.6
Max Probability Value will obviously be the decision or decide the class C1 or C2; which is possible in the test case.
After Normalizing the probabilities:
In this case, P (it will Rain)=0.60
P(it will not Rain)=0.039
Hence ***Decision by Bayesian Classifier: it will rain*** since the P(it will Rain) is higher than the other.
 If any combination of the attribute record and the decision attribute record is not present the table the result might be zero then Laplace Estimator is applied where the numerator is added by one and the Denominator which is the number of records is added by the number of types of class decisions.

 Rational Method
This method is being used to calculate the intensity of the rainfall if it rains on that particular day that is what could be the max or min possible intensity of rainfall it could occur today(based on the month) by using mathematical computations. A database is made which has precipitation records of the particular place recorded since last 10 years, and that will be used to calculate the intensity of rainfall.
 A database is made with month and precipitation attribute so that the required values could be utilized for calculation depending upon the mathematical equation.
 Time of concentration is the important parameter to measure intensity. To determine its 'precise' values numerous way exists but it requires a lot of parameters for which different data is needed and also it depends on the nature of the surface which affects the parameter.

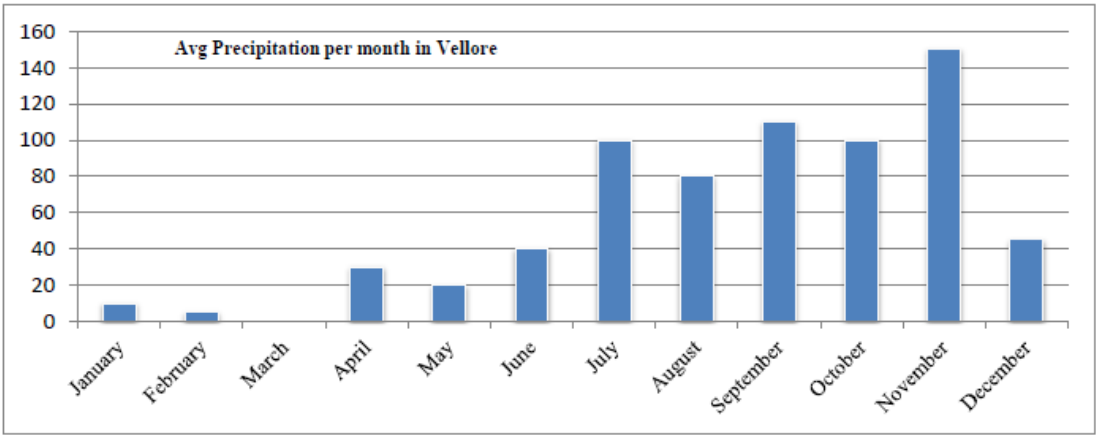
The Test was done in May=> Precipitation observed in Vellore= 20mm (from Figure) (also a separate database) values will be automatically extracted by the code by using ODBC Driver. Precipitation=20mm=20/25=0.8 inch
The month is extracted using the Calendar in Java and then the month is compared with the database month and the corresponding precipitation is extracted using the SQL query.
The precipitation value(in mm) is then converted into inches by dividing the value by 25.4
 The intensity formula is used where the intensity is equivalent to the precipitation which is divided by the time of concentration and the answer SI unit comes in inch/hour.
 The time of concentration, used for the whole process is 5 minutes for max intensity and 60 minutes for min intensity of the rainfall for that particular month.

Hence, Intensity= Precipitation/Tc
For max intensity= (0.8*60)/5=9.69 inch/hr
For min intensity= (0.8*50)/60=0.82 inch/hr
for the month of May.
Tc values have been taken the extremes of the most preferable range.
Hence an Intensity Duration frequency curve can be created by using different values from the range of time of concentrations we are given with. This can give us a pattern of intensity versus time of concentration which can also be useful to observe the rainfall intensity.
The database for the experiment has been prepared for 'Vellore' (Tamil Nadu, India) where the rainfall prediction system was tested.

Hence a database is prepared based on precipitation values every month observed in Vellore for 10 annual years. Below there is a graph between precipitation and months of Vellore region which has been shown which shows axes of precipitation and months. Such precipitation values have been used to create the database table so that the appropriate value could be used for intensity calculation of the rainfall.
Figure 4: Vellore Weather Data [10]

Test Case: (from sensor data & if-else conditions)
Outlook: Sunny
Temperature: Hot
Humidity: Normal
P(Outlook=”Sunny”/C1)= No of times “sunny” occurs in outlook column AND frequency of it will rain from the database=1/3
The Database Rain (Training Data collected for that region “Vellore”):

Figure 3: Training Data



Assumption & Dependencies
Certain assumptions have been taken for the training data itself.
 The Training data is based on assumption such as if the “Temperature” is hot implies it must be greater than 30 and “Humidity” is high it must be greater than 60, hence the ‘outlook’ parameter will be “sunny” as well as the decision will be “No” for rain as it’s a sunny day. So this statement itself is surrounded by two more assumptions that the outlook will be sunny for that particular case and the decision gets influenced by the outlook.

 Three main factors have been considered to make assumptions of temperature ranges (0-10=>cold,10-25=>mild,>25=>Hot) and for humidity (0-40=>Normal,>40=>High) are ***Human Experience, Logical Sense, Online weather reports)***

 The final result of the rain prediction has been, however, is resulted from the Naïve Bayesian Algorithm.
 The time of concentration values has been assumed for the ***preferred range from 5 minutes to 60 minutes*** where the value can be anything in this range or beyond range, but this is the most commonly used range and hence for min and max intensities.

V. Results & Discussions

Table 1: Display of sensor values

1 – Record and Display Temperature and Humidity values on the LCD screen. The system shall record Temperature and Humidity values using appropriate sensors and display these values on the LCD screen as shown below.	
Input	User will connect the Arduino Uno to a laptop using the UART cable and the input will be taken from the Temperature and Humidity Sensors.
Action	The Temperature and Humidity sensors will record the values and the Arduino Code running on the Arduino Open Source Software.
Output	The Temperature and Humidity values are recorded and successfully displayed on the LCD screen display. 
Priority	The very first Step

Table 2: Authentication to the system

2-Webpage Authentication The system shall send the sensor values automatically to the Arduino Code & Webpage	
Input	Along with this the username and password should be manually entered by the user on the web page.
Action	The system will verify the user ID and password in the database. (Optional)
Output	If the User ID and password is verified then the user will get access to the home screen. The Temperature and Humidity values are recorded and successfully displayed on the Webpage. 
Priority	Optional

Table 3: Training Data database

3 – Database Storage The Training data will be stored in the database and appropriate values will be fetched via ODBC driver through the Java Code	
Input	The appropriate values of Outlook, Temperature & Humidity will be set as an input to the system which will be decided by the code based on assumptions and if-else conditions
Action	The input from the system will be used and will be compared according to Naïve-Bayes theorem with the classes to deliver the final result

Output	The Training data database is maintained
	`mysql> use rain;`
	`Database changed`
	`mysql> select * from Rain;`
	`+----------+------+----------+-------+`
	`| outlook | temp | humidity | class |`
	`+----------+------+----------+-------+`
	`| sunny | hot | high | no |`
	`| overcast | cool | normal | yes |`
	`| rainy | cool | normal | yes |`
	`| rainy | mild | high | no |`
	`| sunny | mild | normal | yes |`
	`+----------+------+----------+-------+`
	`5 rows in set (2.23 sec)`

Priority	Essential & Medium priority

Table 4: Result of the Algorithm

4 – Prediction and Display of Result on the Java Program The system shall predict probability of Rainfall using Naïve-Bayes Classifier and then show the result of the prediction on the Java Program	
Input	The database values will be extracted from the database by the Naïve-Bayesian Program(java).
Action	The Naïve-Bayes program will use the Naïve-Bayes Equation to predict results. The required values of Outlook,Temperature & Humidity will be fetched from database via ODBC Driver and probability will be calculated by using the parameters along with the class.

Output	Based on the program probability; rainfall or no rainfall should be displayed and if probability of rainfall is high then the green light will lighten up else the red.
	`C:\Users\Anik\Desktop>javac NaiveBayesian.java`
	`C:\Users\Anik\Desktop>java NaiveBayesian`
	`c1=3`
	`c2=2`
	`total=5`
	`p(c1)=0.6`
	`p(c2)=0.4`
	`Enter Outlook! (sunny/overcast/rainy)`
	`sunny`
	`Enter Temperature:(hot/cool/mild)`
	`hot`
	`Enter Humidity:(normal/high)`
	`normal`
	`Probability that it will rain=0.6056782`
	`Probability that it will not rain=0.039639868`

Priority	Essential & high priority

Table 5: Green LED as a result

5 – Prediction and Display of Result on the Java Program and using the LED lights The system shall predict probability of Rainfall using Naïve-Bayes Classifier and then show the result of the prediction on the Java Program and using LED lights: GREEN for RAIN & RED for NO RAIN.	
Input	The values from the Java program will be entered in the Arduino Program to display the result in LED Lights
Action	If the probability of Rain is higher it will glow GREEN
Output	Based on the program probability of rainfall or no rainfall should be displayed and as the Java program result says that the probability of rainfall is higher hence GREEN LIGHT will lighten up. 
Priority	Essential & high priority

Table 6: Red LED as a Result

6 – Prediction and Display of Result on the Java Program and using the LED lights – MANIPULATED BY LIGHTI A LIGHTER OR MATCHSTICK NEAR THE TEMPERATURE SENSOR.	
Action	A lighter or matchstick will be lighted closer next to the temperature sensor.
Output	Although based on the program probability of rainfall or no rainfall will be displayed but due to manipulation probability of rainfall has been made lower and hence RED LIGHT will lighten up. 
Priority	Demo purpose

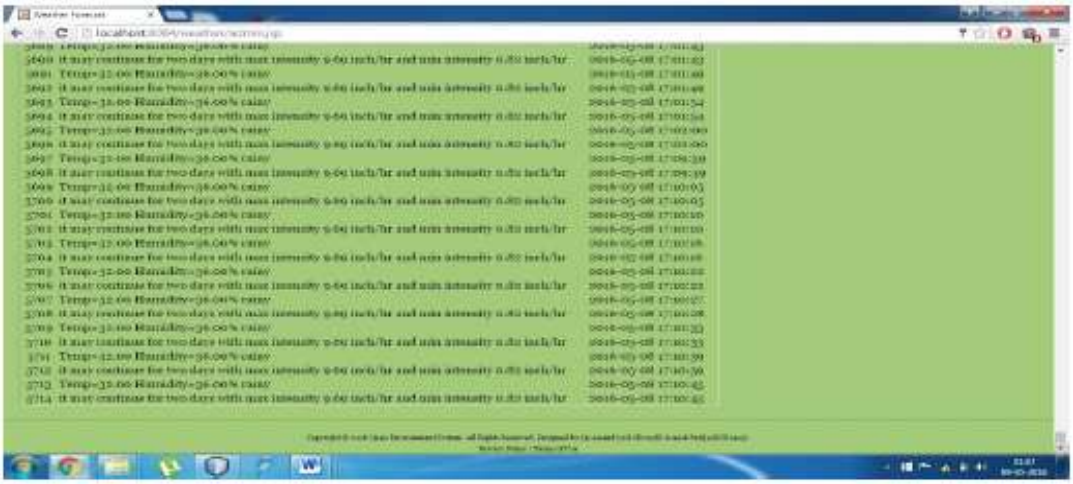
Table 7: Intensity result

7 – Prediction of Intensity The Product shall predict the intensity of Rainfall and reflect in Database as well as the Java Program.	
Input	User will connect the Arduino Uno to a laptop using the UART cable and the input will be taken from the Precipitation Database (Precipitation for the corresponding month). The recorded values will be entered in the Intensity Calculation formula.
Action	The Intensity calculation will use the formula of Precipitation/Tc to give the max and min intensity value.

Output	Based on the program prediction of intensity of rainfall if probability of Rainfall is YES should also be displayed as shown below.
	`C:\Users\Anik\Desktop>javac Intensity.java`
	
	`C:\Users\Anik\Desktop>java Intensity`
	`The Max Intensity of rainfall for ongoing May month will be=9.921259842519685nn`
	`The Min Intensity of rainfall for ongoing May month will be=0.8267716535433072nn`

Priority	Essential & high priority

Figure 8: Final Result on the Webpage



VI. Conclusion

Hence, the rain predicting system predicts the rainfall occurrence along with the intensity of the rainfall if it rains. The temperature and humidity sensor sense the values from the surroundings, and those values are being used to generate the test case for the Naïve Bayesian Classifier in the form of a string. Then based on the training data stored in the database 'rain' the prediction algorithm of Naïve Bayesian Classifier will predict the likelihood of the rainfall or no rainfall on that particular day and whichever value will be greater than the other can be declared as the final result [9]. If it's a sunny day as the final result, we can't expect rain for obvious reasons, but if it's a rainfall day then it will be interesting to know the intensity of rainfall for that particular month. Hence based on the database 'intensity' which stores the value of precipitation levels of that particular area for the ongoing month will be used to calculate the max and min intensity of the rainfall one can expect in that particular month. Hence along with the rainfall, the intensity will be displayed for the particular month and such information can be beneficial for many people to take the necessary steps required based on the information derived from the rain predicting system.

6.1 Scope for future Work

The algorithms will be programmed already for the system, but the user will only have the headache to input the values and rest can be left to the system itself which will perform the necessary actions.

The present system needs to have two separate coding files either in a laptop/PC from where the system can be connected using Power Supply. But the system can be made more user-friendly if the code gets embedded into a chip and can be attached to the portable system. Hence this will obscure the need of separate systems like laptop/PC to carry the code. The assumptions may vary from region to region. Also, we have 3-4 separate coding files for different actions. For example, if-else conditions to generate the test case of outlook, temperature and humidity, there is one file. For Naïve Bayesian programming, we are maintaining a file, for intensity calculation, a small program with a different file is maintained. The main aim & challenge at present can be to integrate everything into a single file. If the required step mentioned above gets fulfilled, the system will get a boost towards quality improvement and will be even more user-friendly and easy to use.

References

Internet Articles:
[1] Tarun Agarwal,Temperature Sensor -Types, Working & Operation, EL-PRO-CUS (The Budding Electronic Engineers Knowledge Space)
[2] SYHITECH SY-HS-220, Humidity Sensors, Electronic components Journal Papers:
[3] Ivan Simeonov, Hristo Killifarev and Raycho Ilarionov, Embedded System for Short Term Weather Forecasting", International Conference on Computer Systems and Technologies,2006.
News Articles:
[4] Express News Service, Agency makes model to forecast for 20 days in advance every 5 days, Indianexpress.com,2015.
[5] Sakaal Times,New Weather forecasting system launched, Sakaaltimes.com,2013.
Book Chapters:
[6] Bjerknes Kyrner,Weather Prediction by Numeral process,Chapter-1,1996.
[7] Nautral Resources Management and Environment Department, Rainfall and Evapotranspiration, FAO Corporate Document Repository, Chapter-4
[8] NPTEL, Module 6-Quantity Estimation of Storm Water,IIT Kharagpur Web Courses.
Internet Material:
[9] Naïve Bayesian Classification Algorithm,software.ucv.ro, AIR Doc 4.
[10] World Weather Online, Vellore,Tamil Nadu Monthly Climate Average, India.

IOSR Journal of Computer Engineering (IOSR-JCE) is UGC approved Journal with Sl. No. 5019, Journal no. 49102.

Anik Pait. "The Open Environment System: Rainfall Prediction Using Naïve Bayesian Algorithm." IOSR Journal of Computer Engineering (IOSR-JCE), vol. 19, no. 5, 2017, pp. 01–10.