# VISVESVARAYA TECHNOLOGICAL UNIVERSITY

**'JNANA SANGAMA' BELAGAVI-590 018, KARNATAKA**

**PROJECT REPORT**

**ON**

## "TEMPERATURE AND HUMIDITY PREDICTION USING MACHINE LEARNING"

**SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE,**

**BACHELOR OF ENGINEERING**
**IN**
**COMPUTER SCIENCE & ENGINEERING**

**Submitted By**


1. **MANOJ KUMAR D** [1CG17CS053]
2. **NANDAN KUMAR T N** [1CG17CS060]
3. **KIRAN M** [1CG17CS044]
4. **DARSHAN H N** [1CG17CS020]


**Under the guidance of:**


**Mrs. Vidhya H A** M.Tech.,
Asst. Prof., Dept. of CSE,
CIT, Gubbi, Tumakuru.


**HOD:**


**Dr. Shantala C P**, Ph.D
Prof & Head, Dept. of CSE,
CIT, Gubbi, Tumakuru.



**Channabasaveshwara Institute of Technology**
**(NAAC Accredited & ISO 9001:2015 Certified Institution)**
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.


**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2020-21**

**Channabasaveshwara Institute of Technology**
**(NAAC Accredited & ISO 9001:2015 Certified Institution)**
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.
**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2020-21**

## DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

## CERTIFICATE

This is to certify that the project work entitled **"Temperature and Humidity Prediction Using Machine Learning"** has been successfully carried out by **Manoj Kumar D [1CG17CS053], Nandan Kumar T N [1CG17CS060], Kiran M [1CG17CS044], Darshan H N [1CG17CS020],** bonafide students of **CHANNABASAVESHWARA INSTITUTE OF TECHNOLOGY, GUBBI, TUMAKURU,** under our supervision and guidance and submitted in partial fulfillment of the requirements for the award of Degree in **Bachelor of Engineering** by **Visvesvaraya Technological University, Belagavi** during the academic year of 2020–2021. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements for the above said degree.

**Guide:**

**Mrs. Vidya H A** M.Tech.,
Asst. Prof., Dept. of CSE,
CIT, Gubbi, Tumakuru.

**H.O.D:**

**Dr. Shantala C P**, Ph.D
Prof & Head, Dept. of CSE,
CIT, Gubbi,Tumakuru.

**Principal:**

**Dr. Suresh D S** Ph.D
CIT, Gubbi, Tumakuru.

**Examiners:**

1.

2.

**Channabasaveshwara Institute of Technology**
**(NAAC Accredited & ISO 9001:2015 Certified Institution)**
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.

**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2020-21**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

# UNDERTAKING

We the students **Manoj Kumar D [1CG17CS053], Nandan Kumar T N [1CG17CS060], Kiran M [1CG17CS044], Darshan H N [1CG17CS020]** of **VIII semester B.E. Computer Science and Engineering** of **CHANNABASAVESHWARA INSTITUTE OF TECHNOLOGY, GUBBI, TUMAKURU** declare that Project work entitled **"Temperature and Humidity Prediction Using Machine Learning"** has been carried out and submitted in partial fulfillment of the requirements for the award of degree in Bachelor of Engineering in **Computer Science and Engineering** by the Visvesvaraya Technological University during the academic year 2020-2021.

1. **Manoj Kumar D [1CG17CS053]**
2. **Nandan Kumar T N [1CG17CS060]**
3. **Kiran M [1CG17CS044]**
4. **Darshan H N [1CG17CS020]**

**Channabasaveshwara Institute of Technology**
**(NAAC Accredited & ISO 9001:2015 Certified Institution)**
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.

**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2020-21**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

# BONAFIDE CERTIFICATE

This is to certify that the Project work entitled **"Temperature and Humidity Prediction Using Machine Learning"** is a bonafide work of **Manoj Kumar D [1CG17CS053], Nandan Kumar T N[1CG17CS060], Kiran M [1CG17CS044], Darshan H N [1CG17CS020]** students of **VIII** semester B.E computer Science and Engineering carried out at **Channabasaveshwara Institute of Technology**, Gubbi, Tumakuru, in partial fulfillment of the requirements of the award of degree in B.E. in **Computer Science and Engineering** of Visvesvaraya Technological University, Belagavi under my supervision and guidance. Certified that to the best of my knowledge the work reported here in does not from part of any other thesis on the basis of which degree or award was conferred on earlier occasion to this or any other candidates.

**Guide:**

**Mrs. Vidya H A** M.Tech.,
**Asst. Prof., Dept. of CSE,**
**CIT, Gubbi, Tumakuru.**

# ACKNOWLEDGEMENT

A great deal of time and lot of effort has gone into completing this project report and documenting it. The number of hours spent in getting through various books and other materials related to this topic chosen by us have reaffirmed its power and utility in doing this project.

Several special people have contributed significantly to this effort. First of all, we are grateful to our institution "**Channabasaveshwara Institute of Technology"**, Gubbi which provided us an opportunity in fulfilling our most cherished desire of reaching the goal.

We acknowledge and express our sincere thanks to the beloved Director and Principal **Dr. Suresh D S** for his many valuable suggestions and continued encouragement and support in the academic endeavors.

We wish to express our deep sense of gratitude to **Dr. Shantala C P,** Head of the Department of Computer Science and Engineering for all the guidance and who still remains a constant driving force and motivated through innovative ideas with tireless support and advice during the project to examine and helpful suggestions offered.

We wish to express our deep sense of gratitude to our guide **Mrs. Vidya H A, M.Tech** Dept of CSE, CIT, for his meticulous attention to details, which has contributed immeasurably to the quality of the project report.

This would never been possible without the support and technical supervision by all the faculty members of CITRIS for all their guidance.

We would express our gratitude towards our parents and friends for their kind cooperation and encouragement which helped us in completion of this project.

Finally, we would like to thank all the teaching and non-teaching staff of Dept of CSE, for their cooperation.

Thanking everyone….

**Manoj Kumar D [1CG17CG17CS053]**

**Nandan Kumar T N [1CG17CS60]**

**Kiran M[1CG17CS044]**

**Darshan H N [1CG17CS020]**

# ABSTRACT

Temperature and Humidity prediction is the application of science and technology to predict the state of the atmosphere for a given location. Ancient weather forecasting methods usually relied on observed patterns of events, also termed pattern recognition. For example, it might be observed that if the sunset was particularly red, the following day often brought fair weather. However, not all of these predictions prove reliable. Here this system will predict weather based on parameters such as temperature, humidity and wind. This system is a web application with effective graphical user interface. User will login to the system using his user ID and password. User will enter current temperature; humidity and wind, System will take this parameter and will predict weather from previous data in database. The role of the admin is to add previous weather data in database, so that system will calculate weather based on these data. Weather forecasting system takes parameters such as temperature, humidity, and wind and will forecast weather based on previous record therefore this prediction will prove reliable. This system can be used in Air Traffic, Marine, Agriculture, Forestry, Military, and Navy etc.

# CONTENTS

# CHAPTER 1

## INTRODUCTION

Synoptic data or climate data are the two classifications of weather data. Synoptic data is the real-time data provided for use in aviation safety and forecast modelling. Climate data is the official data record, usually provided after some quality control is performed on it. Climate and weather affect the human society in all the possible ways. Crop production in agriculture, the most important factor for water resources i.e., Rain, an element of weather, and the proportion of these elements increases or decreases due to change in climate. Energy sources, e.g. natural gas and electricity are greatly depends on weather conditions. Climate is not fixed, the fluctuation in the climate can be seen from year to year, e.g., rain/dry; cold/warm seasons significantly influence society as in all the possible ways. Depending upon the techniques used Data Mining can be divided into three basic types, i.e., Association Rules Mining, Cluster analysis and Classification/Prediction.

Weather forecasting system takes parameters such as temperature, humidity, and wind and will forecast weather based on previous record therefore this prediction will prove reliable. This system can be used in Air Traffic, Marine, Agriculture, Forestry, Military, and Navy etc. Forecasting the temperature and rain on a particular day and date is the main aim of this paper. In the paper we forecast rain and temperature for Europe; year up to 2051 and also, we forecast temperature of world; year up to 2100.Our paper is aimed to provide real time weather forecast service at finest granularity level with recommendations. We grab user's location (longitude, latitude) using GPS data service whenever user requests for our services. Our system will process the users query and will mine the data from our repository to draw appropriate results. Users will be provided with recommendations also and that is the key facility of our service. Personalized forecast is generated for each individual user based on their location.

## 1.1 Problem statement:

Heavy rainfall can lead to numerous hazards, for instance:

flooding, including danger to human life, harm to structures and framework, and loss of products and domesticated animals. avalanches, which can compromise human life, upset transport and interchanges, and cause harm to structures and foundation. Where overwhelming precipitation happens with high breezes, hazard to ranger service crops is high.

For example, if we consider an area affected by tropical cyclone the fundamental impacts of tropical cyclone incorporate heavy rain, strong wind, huge tempest floods close landfall, and tornadoes. The devastation from a tropical cyclone, for example, a sea tempest or hurricane, depends for the most part on its power, its size, and its area. Tropical tornados act to evacuate woods shade and additionally change the scene close beach front zones, by moving and reshaping sand ridges and causing broad disintegration along the drift. Indeed, even well inland, overwhelming precipitation can prompt mudslides and avalanches in rugged regions. Their belongings can be detected after some time by concentrate the convergence of the Oxygen-18 isotope inside caverns inside the region of typhoons' ways. So, we are providing a better way to get accurate predictions. As mentioned above, the benefits of identifying important features of mechanical learning, complex data sets, play an important role in forecasting of weather. Since the best results can be achieved with engineering learning algorithms, we should use these techniques to aware people from natural disasters. This is because learning engineering algorithms can provide more accurate results. Apart from this, the results are achieved at a short time and people get enough time to do preparations or to escape from that place.

## 1.2 Objective:

Accurate temperature and Humidity prediction helps to safeguard life and property, playing an important role in planning activities for the government, industry, and the public. The primary aim of this project is to review the different machine learning strategies for temperature forecasting. This survey shows that Machine Learning techniques can help to accurately predict temperatures based on a set of input features, which can include the previous values of temperature, relative humidity, solar radiation, rain and wind speed measurements, among others.

## 1.3 Future Scope:

We discovered that feature scaling is an essential aspect of ML models during this project. The basic concept is to make sure that all of the functionality are on the same scale. We're just trying to get things moving faster here; the aim is to get all of the input variables into one of these ranges, give or take a few.

In the coming years, we should strive to reduce variance as much as possible, as this will aid in better prediction, resulting in a successful ML model. Since outliers are bad for ML models, they should be avoided or omitted before determining the best match. This not only improves the model's precision, but it also keeps the findings consistent, which could be different if outliers were included.

Professional weather forecasters aren't flawless, but they're usually more reliable than this linear regression model's predictions. Weather is a non-linear system, according to this. Furthermore, unlike most forecasters, I based all of my predictions on data from a single location rather than several locations.

# CHAPTER 2

## LITERATURE SURVEY

Prediction of the future values by analysing Temperature and humidity data is one of the important parts which can be helpful to the society as well as to the economy. Work has been done in this constrain since years. Different techniques have been applied to predict the temperature and humidity and other parameters of weather. Some of the work in this area is as follows: In data mining, the unsupervised learning technique of clustering is a useful method for ascertaining trends and patterns in data. Most general clustering techniques do not take into consideration the time-order of data. Tasha R. Innis used a mathematical programming and statistical techniques and methodologies to develop a seasonal clustering technique for determining clusters of time series data, and applied this technique to weather and aviation data to determine probabilistic distributions of arrival capacity scenarios, which can be used for efficient traffic flow management.

The seasonal clustering technique is modelled as a set partitioning integer programming problem and resulting clustering's are evaluated using the mean square ratio criterion. The resulting seasonal distributions, which have satisfied the mean square ratio criterion, can be used for the required inputs (distributions of airport arrival capacity scenarios) into stochastic ground holding models. In combination, the results would give the optimal number of flights to ground in a ground delay program to aid more efficient traffic flow management. S. Kotsiantis, A. Kostoulas, S. Lykoudis, A. Argiriou, K. Menagias proposed a hybrid data mining technique that can be used to predict more accurately the mean daily temperature values , it was found that the regression algorithms could enable experts to predict temperature values with satisfying accuracy using as input the temperatures of the previous years. The hybrid data mining technique produce the most accurate results.

# CHAPTER 3

## METHODOLGY

There are two primary phases in the system:

1. Training phase: The system is trained by using the data in the data set and fits a model (line/curve) based on the algorithm chosen accordingly.

2. Testing phase: the system is provided with the inputs and is tested for its working. The accuracy is checked. And therefore, the data that is used to train the model or test it, has to be appropriate.

The system is designed to detect and predict price of used car and hence appropriate algorithms must be used to do the two different tasks. Before the algorithms are selected for further use, different algorithms were compared for its accuracy. The well-suited one for the task was chosen.

We have a problem statement which comes under the category of Classification. It is a multiclass classification in which the classes given to us are

1. Rain, tempest, and snow into precipitation

2. For the most part shady, foggy, and cloudy into exceptionally shady

3. Scattered mists and somewhat shady into modestly shady

4. Clear as clear Our aim is to classify the given data into the above given classes.

In order to do so, we have to first analyse the data given to us. For analysing the features, we are using different techniques. The training of model can be done in many ways. It depends on how the data is prepared for further processing. The data can be used directly depending on the situation or the data can be used to form a histogram. After these modifications, we choose a particular model on which we will train our data. This model can be: Linear regression, Logistic Regression, SVM, Neural Networks, Decision Tress, K-Nearest Neighbors etc. Parameter tuning can also be done in order to increase our accuracy.

Once the model is trained, we can test our data by applying our algorithms on the Test Data. With the help of this we can find the learning ability of our algorithm.

# CHAPTER 4

## SYSTEM DESIGN:

The record has just been separated into train set and test set. Each information has just been labelled. First, we take the trainset organizer.

We will train our model with the help of histograms. The feature so extracted is stored in a histogram. This process is done for every data in the train set. Now we will build the model of our classifiers. The classifiers which we will take into account are Support Vector Machines, and Artificial Neural Networks. With the help of our histogram, we will train our model. The most important thing to in this process is to tune the parameters the accordingly, such that we get the most accurate results. Once the training is complete, we will take the test set. Now for each data variable of test set, we will extract the features using feature extraction techniques and then compare its values with the values present in the histogram formed by train set. The output is then predicted for each test day. Now in order to calculate accuracy, we will compare the predicted value with the labelled value. The different metrics that we will use are confusion matrix, accuracy score, f1 score etc.

### 4.1 EXISTING SYSTEM:

Preparation of All India weather bulletins that includes forecast & warning with graphics along with text up to 3 days for 36 sub-divisions and outlook for subsequent 4 days. It has been issued four times in a day and disseminated to various users like Door Darshan, All India Radio, Press Information Bureau, National Disaster Management Authority, National Disaster Management, Ministry of Home affairs, Ministry of Agriculture etc. This is also post in IMD Website for public in general.

a. Preparation of 5 days forecast of different cities at All India Level, to preparation of this forecast current synoptic observations, satellite imageries and numerical Temperature and Humidity Prediction model outputs taken into consideration.

b. Preparation of Press Releases/Reports on severe weather events. During monsoon season, weekly press release with detail weather of the week and forecast

& warning for one week.

c. Monitoring of southwest and northeast monsoon onset, advance, performance and its withdrawal.

d. Co-ordination with sub offices of IMD regarding forecast & warning through video/audio conferences.

e. Collection of meteorological parameters like temperature, rainfall, snowfall, humidity etc. of major cities for internal use and for media briefing etc.

f. To provide the operational forecasting training to national and international forecasters. Also impart internship training to various colleges and university students. In addition, also provide briefing to school/college/university students, disaster managers and other various national and international delegates.

## 4.2 PROPOSED SYSTEM:

Our model collects historical weather data that includes a variety of important factors that influence weather change, such as temperature, both maximum and minimum temperatures, atmospheric moisture or humidity, precipitation, the UV Index of the atmosphere, and atmospheric mean pressure. In our proposed model, the collected dataset is divided into sections that are useful to the machine learning model and parts that aren't. After that, the dataset goes through data pre-processing, which involves passing the data through a process that replaces missing and error values in the dataset with mean values or the most frequently occurring value in that field.

Following the data pre-processing, the cleaned dataset is divided into two parts: the training set and the test set. The training set is used to teach the machine learning model how to compute the results, while the testing set is used to find the results, compare the real and measured values, and use the error value as a benchmark to teach the machine learning model even more.

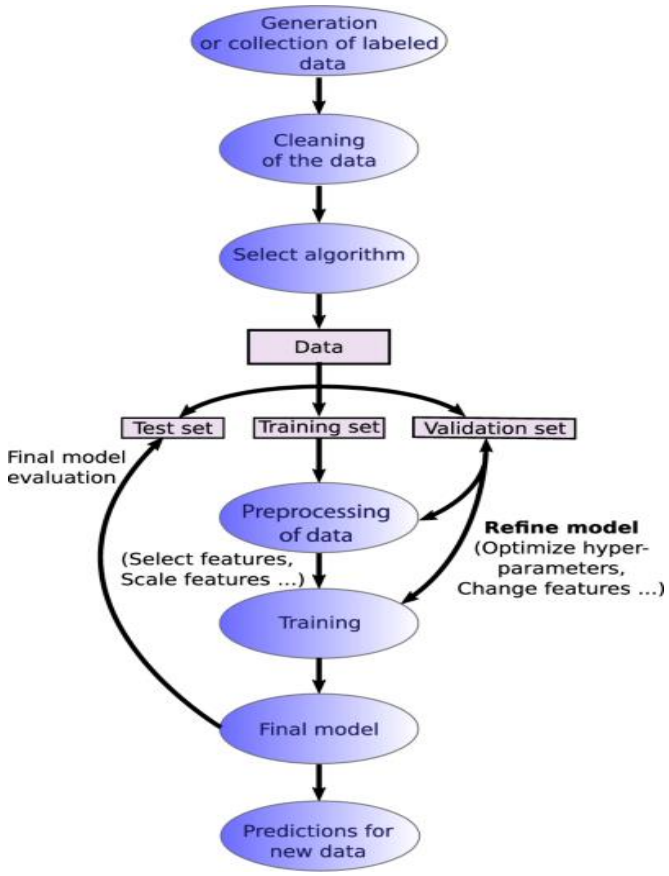



**Fig – 1: Proposed System Flowchart**

## 4.3 LIBRARIES USED:

• NumPy (for Numerical Analysis)

• Pandas (for handling data files)

• Matplotlib (for visualizations inline & figure settings)

• Seaborn (for better relational visualizations)

• Scikit Learn (for model building & data pre-processing)

## 4.4 TECHNOLOGY USED:

### MACHINE LEARNING:

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so.

Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

## 4.5 DATASET:

For this project, we are using the dataset on used car sales from all over the United States, available on Kaggle.

Depending on the location of the weather stations, it offers current as well as historical weather data going back many years. The data set given by the source is fairly accurate and includes a variety of features to work with, such as mean temperature, maximum temperature, humidity, and precipitation. The dataset has 619999 rows and four features.

| | id | temp | humidity | date_time |
|---|---|---|---|---|
| **0** | 9197 | 21 | 20 | 2016-08-17 00:00:01 |
| **1** | 9198 | 21 | 20 | 2016-08-17 00:01:01 |
| **2** | 9199 | 21 | 20 | 2016-08-17 00:02:01 |
| **3** | 9200 | 21 | 20 | 2016-08-17 00:03:01 |
| **4** | 9201 | 21 | 20 | 2016-08-17 00:04:01 |
| **...** | ... | ... | ... | ... |
| **619994** | 629191 | 18 | 21 | 2017-12-17 03:13:01 |
| **619995** | 629192 | 18 | 21 | 2017-12-17 03:14:02 |
| **619996** | 629193 | 18 | 21 | 2017-12-17 03:15:02 |
| **619997** | 629194 | 19 | 20 | 2017-12-17 03:16:01 |
| **619998** | 629195 | 19 | 20 | 2017-12-17 03:17:02 |

619999 rows × 4 columns

## 4.6 ADVANTAGES:

This is useful if we have unanticipated values to consider. We should pay attention to whether the value is very high or very low.

## 4.7 DISADVANTAGES:

If we make a single extremely bad prediction, squaring may amplify the error and can skew the metric to overestimate the model's badness. That is an especially problematic behaviour when we have noisy data (data that is not completely accurate for any reason) — even a perfect model may have a high MSE in that case, making it difficult to judge how well the model is doing. On the other hand, if all of the errors are small, or even smaller than one, the opposite effect occurs: we can underestimate the model's poor performance.

## 4.8 ALGORITHM USED:

### Linear Regression:

Linear Regression attempt to model the relationship between two variables by fitting a linear equation to observed data. The other is considered to be dependent variable. For Example: A modeller might want to relate weights of individuals to their heights using a linear regression model.

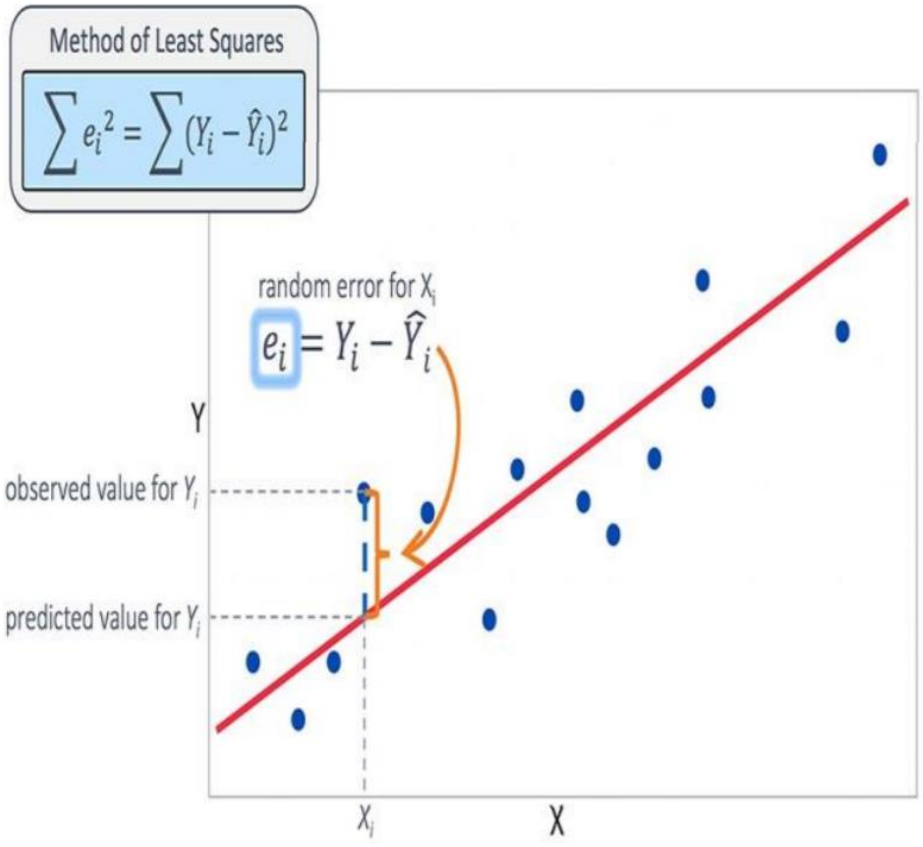



**Fig – 2: Linear Regression**

Linear regression is useful for finding relationship between multiple continuous variables. There are multiple independent variables and single independent variable

$y = m1X1+m2X2+\ldots\ldots+b$

m1, m2, m3 …. →slope

b → y intercept

X1, X2, X3 …… →independent variables

y → dependent variables.

As the function is a regression model, score function will help us find the accuracy of our model. Our model will be much accurate when the score is nearer to 1.0

## K-Means Algorithm:

k-means is one of the most straightforward unsupervised learning calculations that take care of the notable bunching issue. The methodology pursues a basic and simple approach to group a given informational collection through a specific number of bunches (expect k bunches) fixed apriority. The primary thought is to characterize k focuses, one for each bunch. These focuses ought to be put in a cleverness route in light of various area causes distinctive outcome. In this way, the better decision is to put them however much as could be expected far from one another. The subsequent stage is to take each guide having a place toward a given informational collection and partner it to the closest focus.

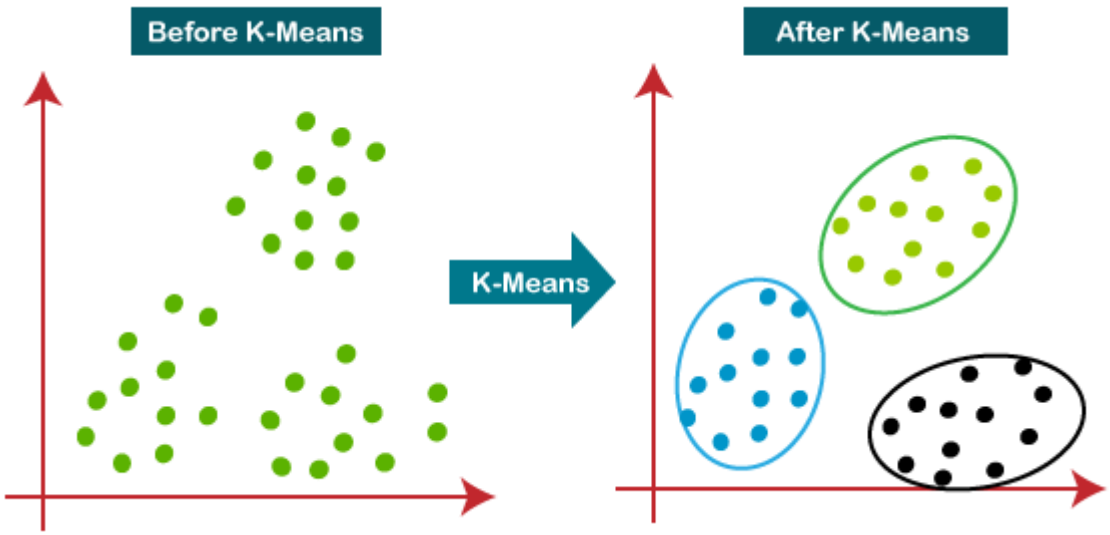



**Fig – 2: K-Means Algorithm**

Algorithm:

**Algorithm 1** $k$-means algorithm

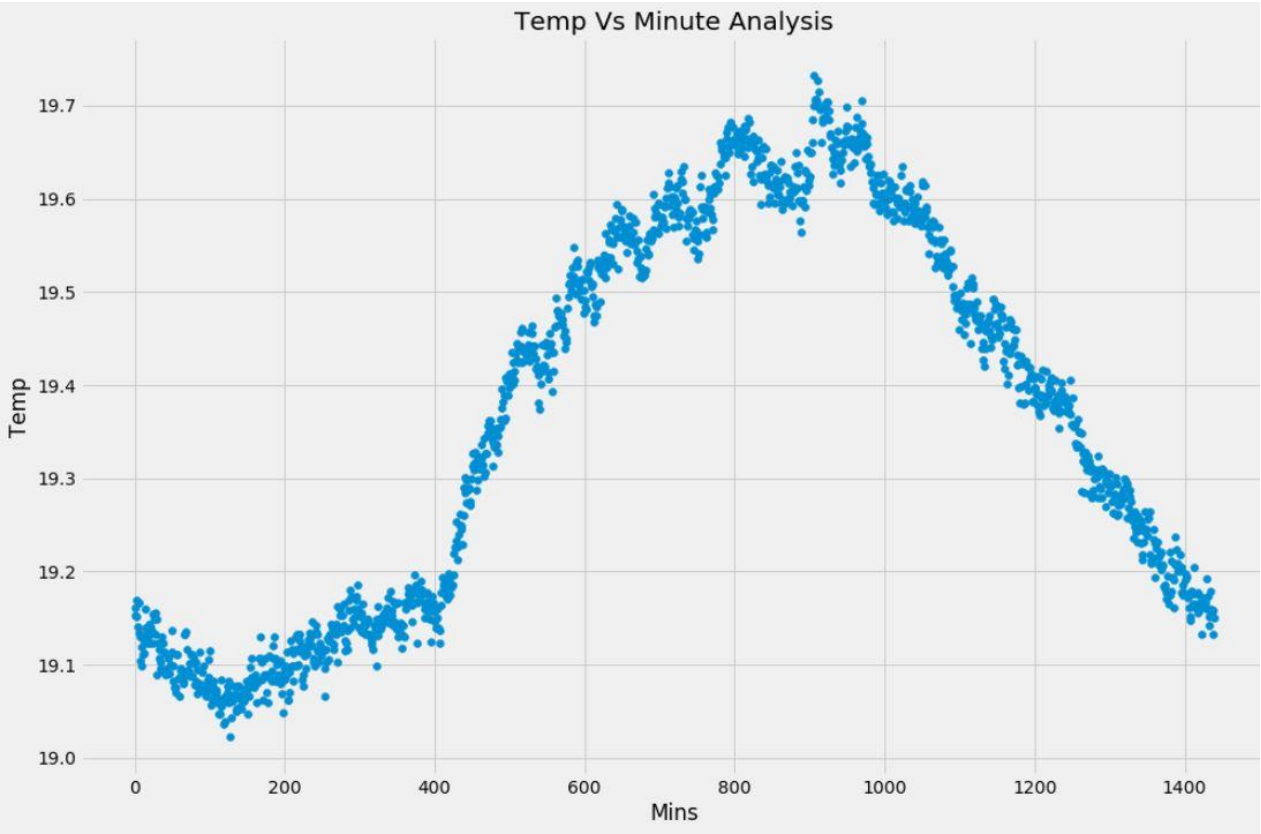
1: Specify the number $k$ of clusters to assign.
2: Randomly initialize $k$ centroids.
3: **repeat**
4: **expectation:** Assign each point to its closest centroid.
5: **maximization:** Compute the new centroid (mean) of each cluster.
6: **until** The centroid positions do not change.

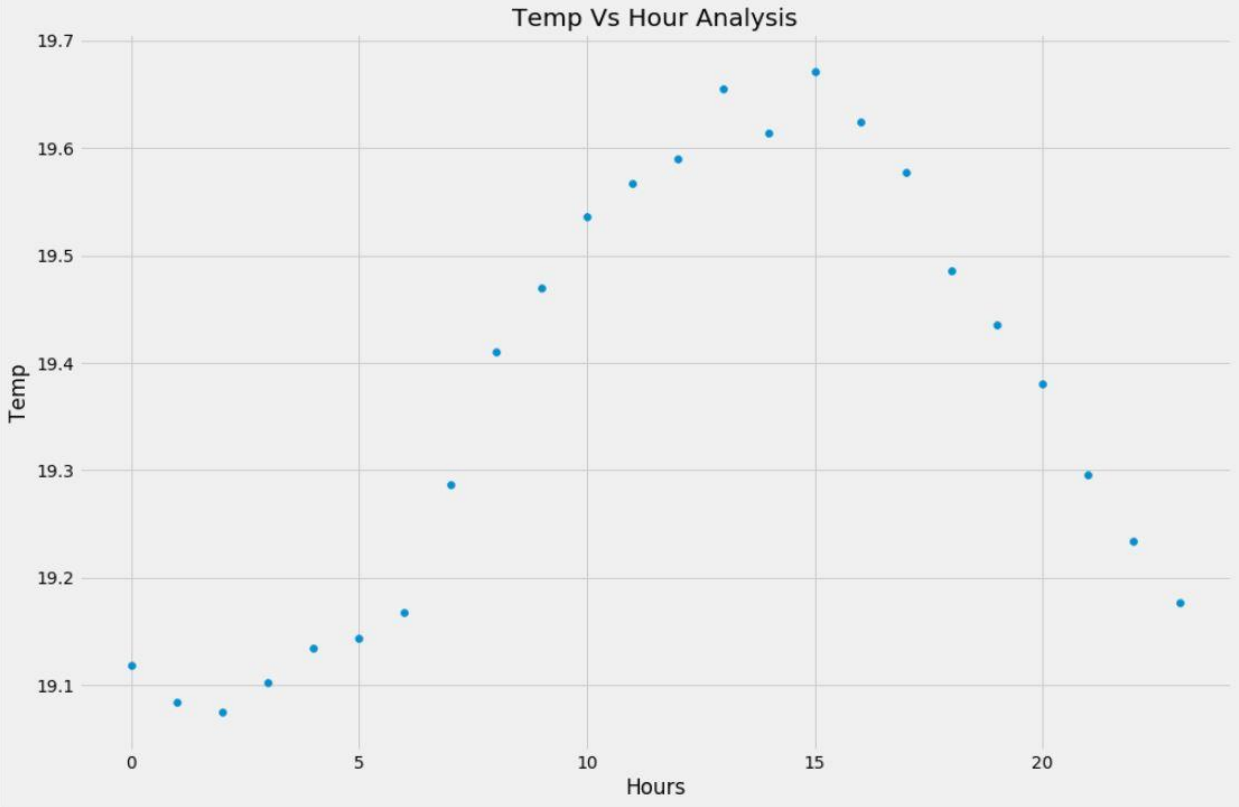
# CHAPTER 5

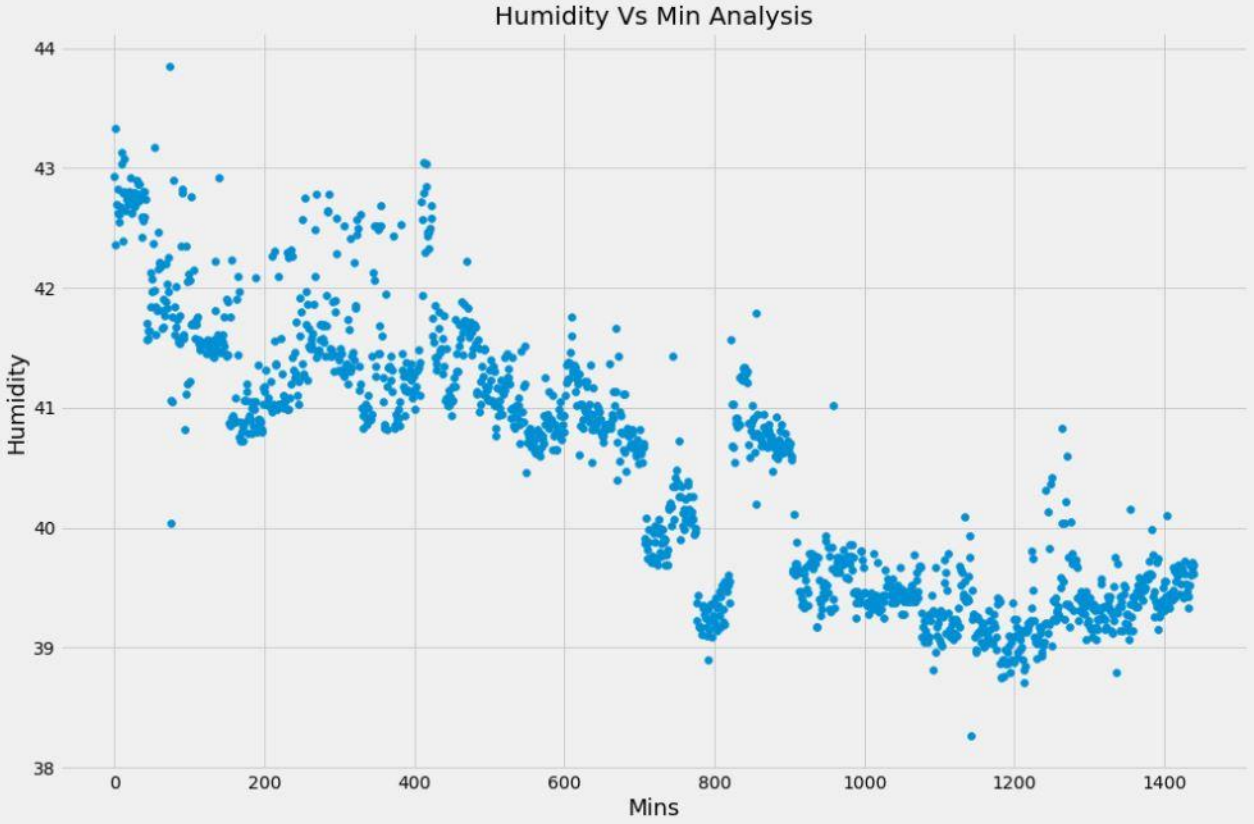
## IMPLEMENTATION

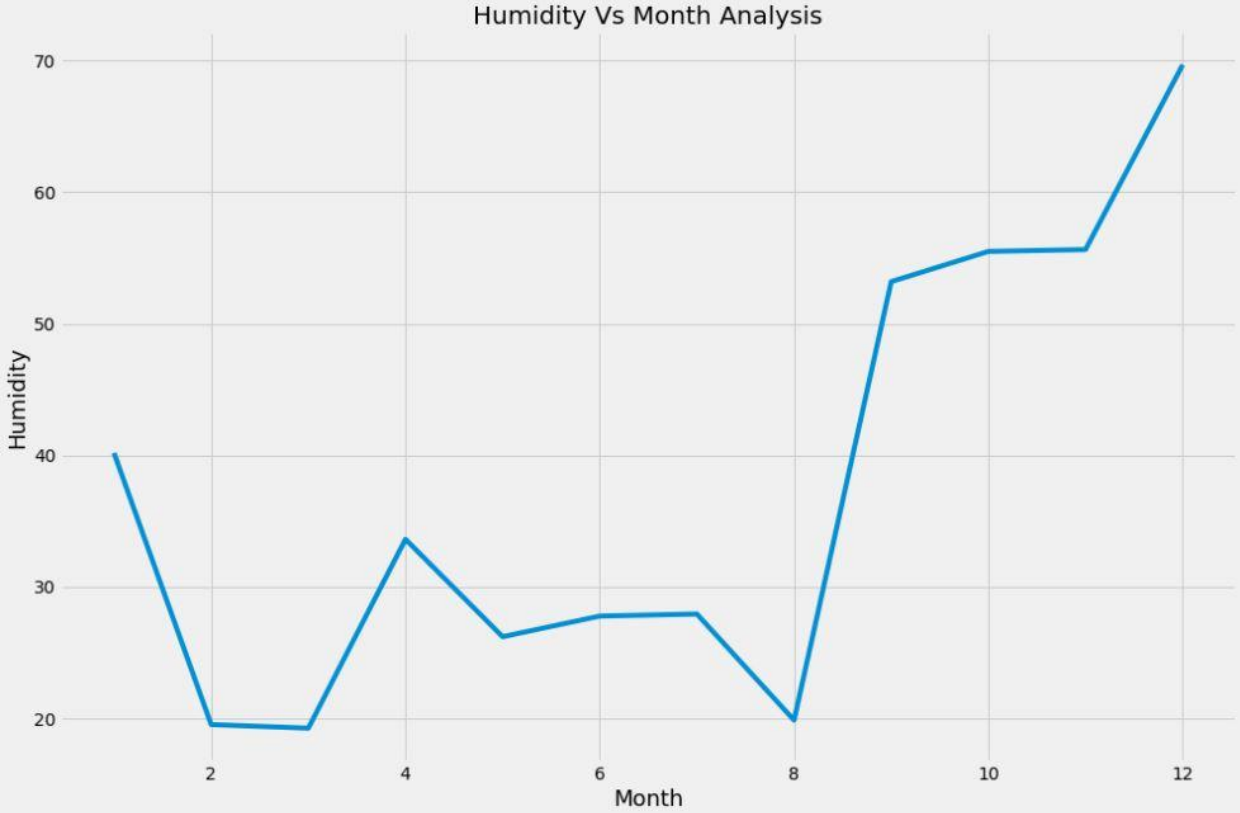
To be used efficiently, all computer software needs certain hardware components or other software resources to be present on a computer. These prerequisites are known as (computer) system requirements and are often used as a guideline as opposed to an absolute rule. Most software defines two sets of system requirements: minimum and recommended. With increasing demand for higher processing power and resources in newer versions of software, system requirements tend to increase over time. Industry analysts suggest that this trend plays a bigger part in driving upgrades to existing computer systems than technological advancements.

## 5.1 DATA VISUALIZATION:




Temp Vs Hour Analysis
Temp
Hours
19.7
19.6
19.5
19.4
19.3
19.2
19.1
0
5
10
15
20

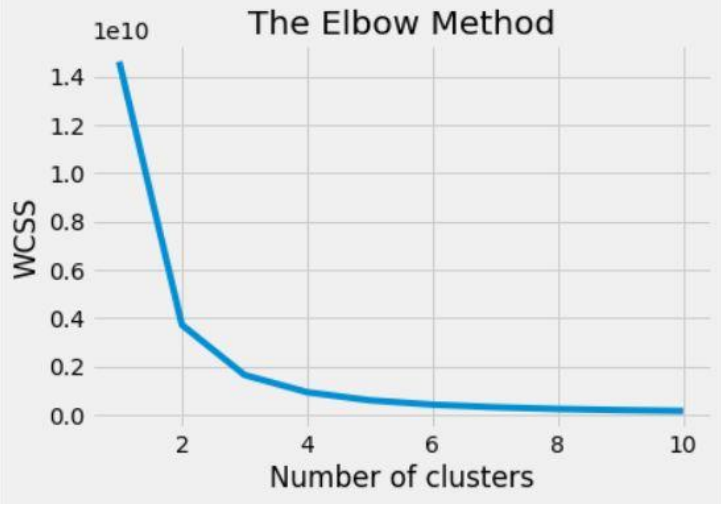

Humidity Vs Min Analysis
Humidity
Mins
44
43
42
41
40
39
38
0
200
400
600
800
1000
1200
1400

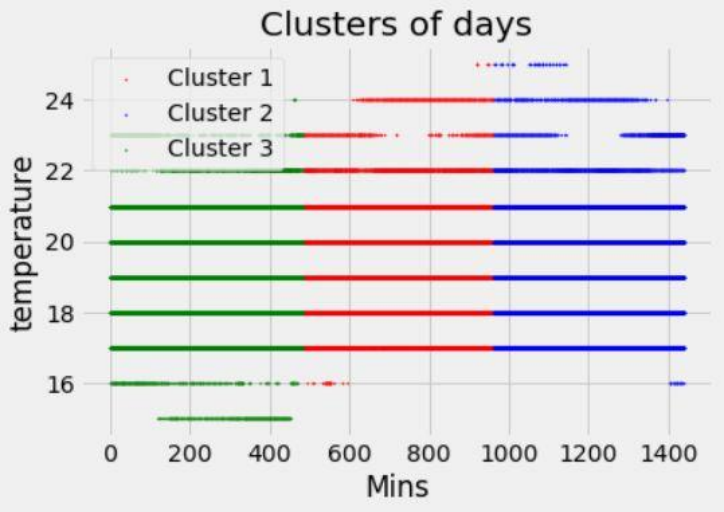



## 5.2 MODEL BUILDING:

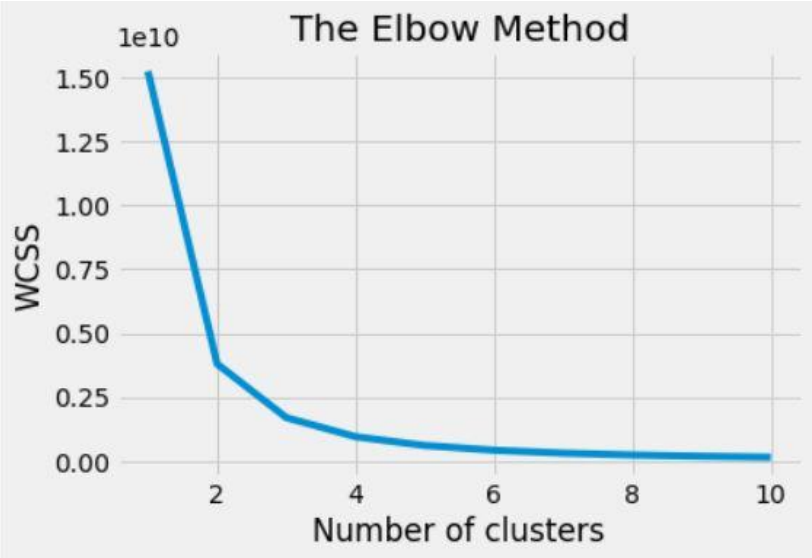
We used K-means Classification algorithm to build a model. We divided the data into seven parts, which can be classified into day wise data. Elbow method is uesd to find the number of clusters and find mean error.

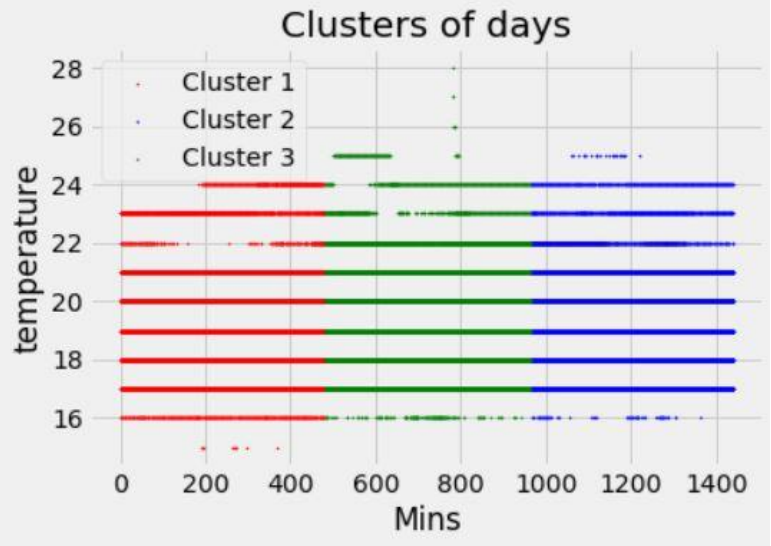
Elbow method and clusterung for **Monday**:



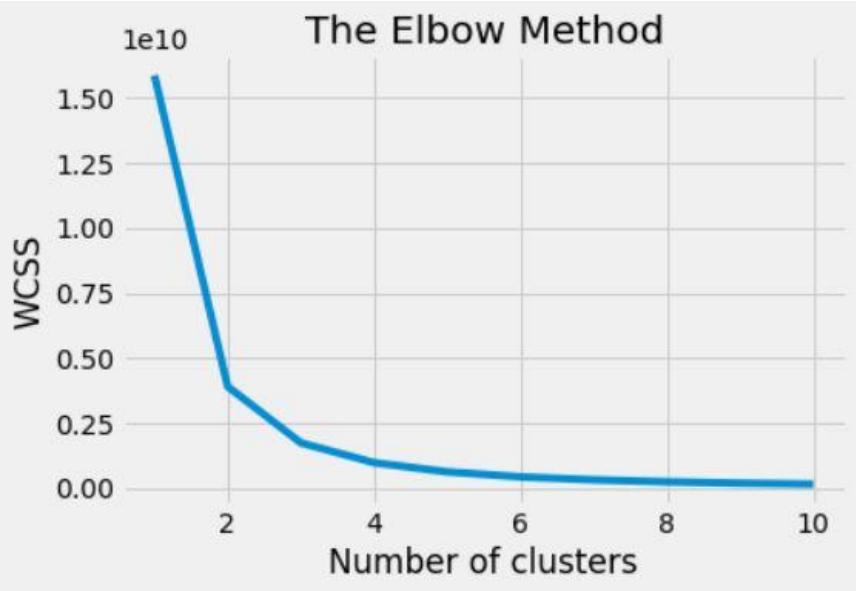


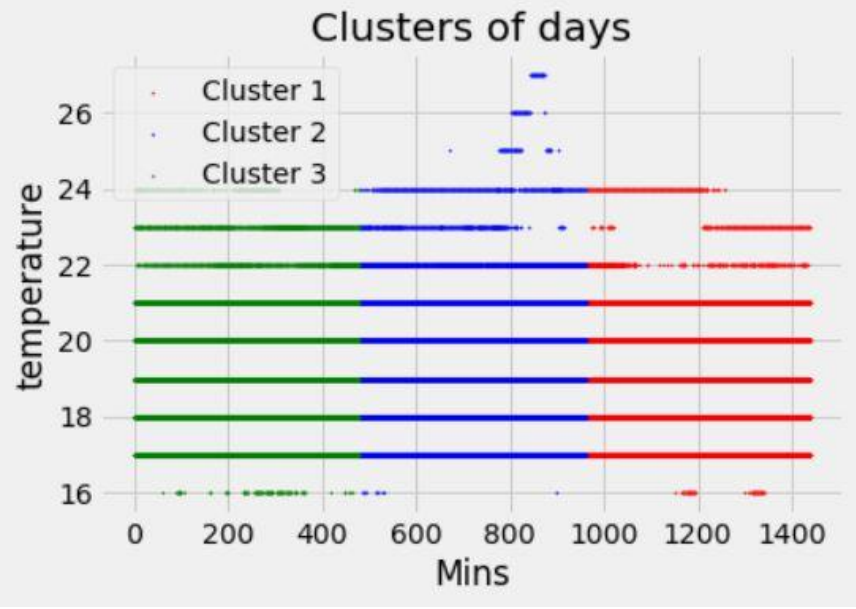

Elbow method and clusterung for **Tuesday**:



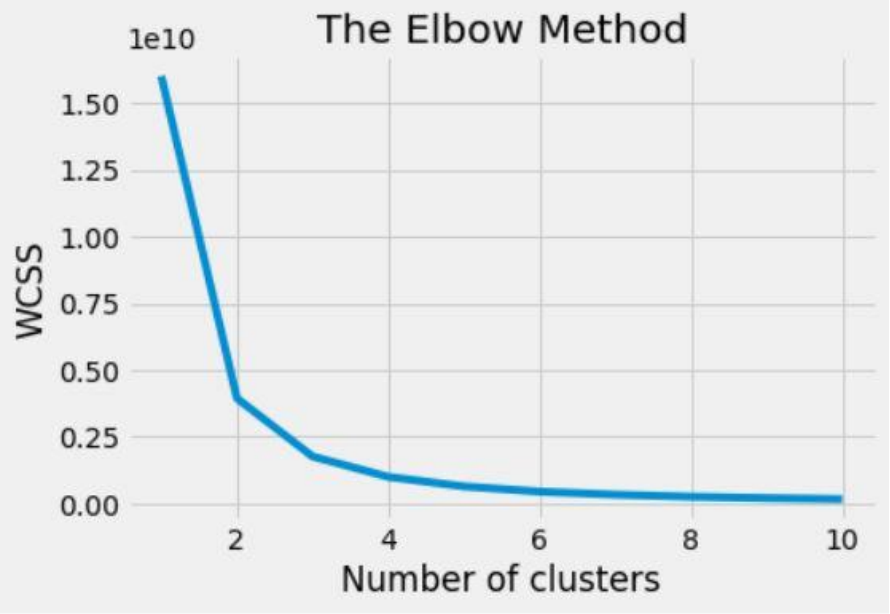



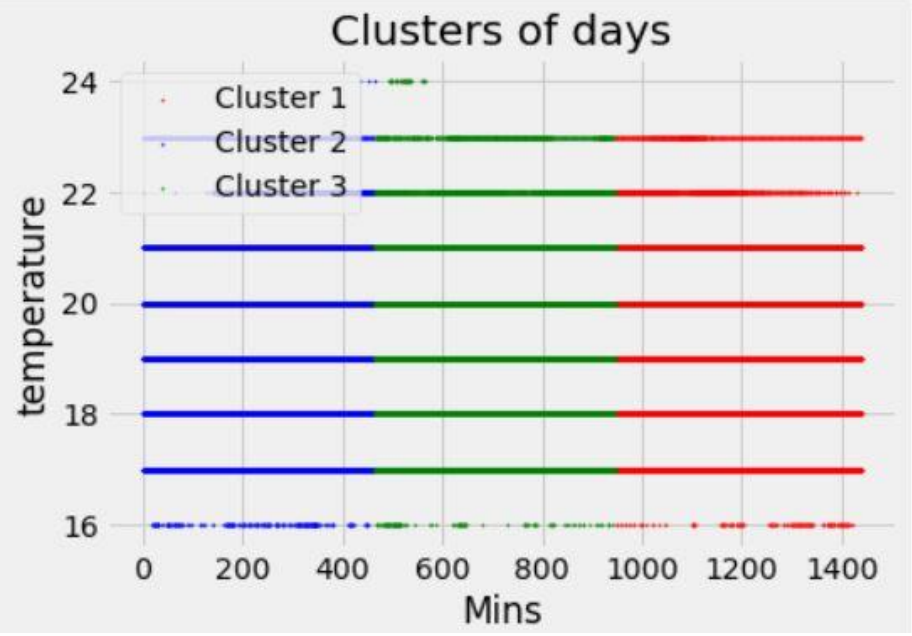
Elbow method and clusterung for **Wednesday**:



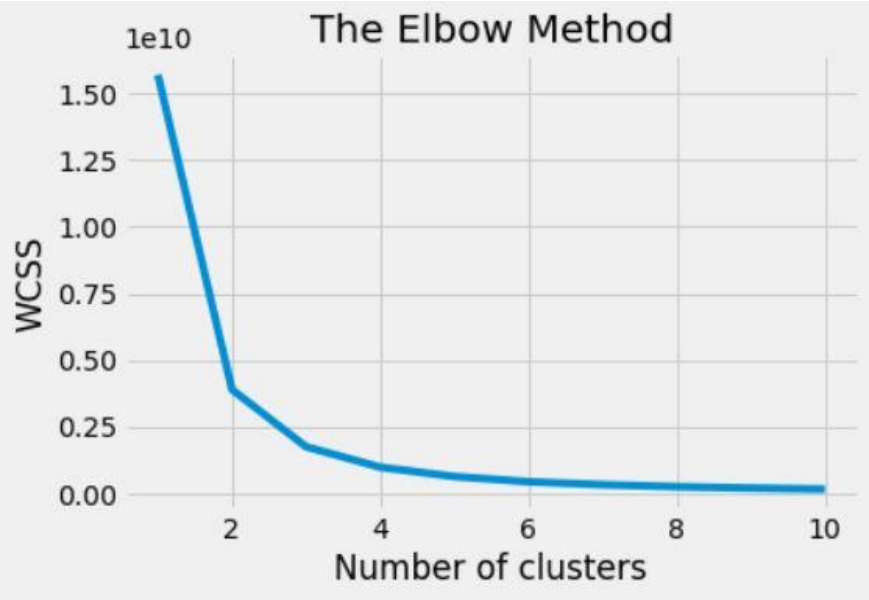



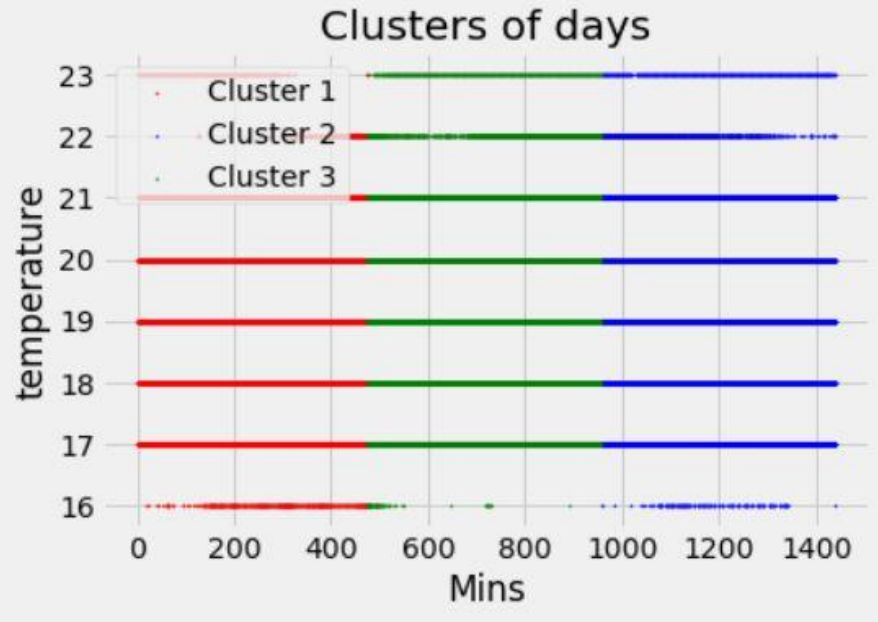
Elbow method and clusterung for **Thursday**:



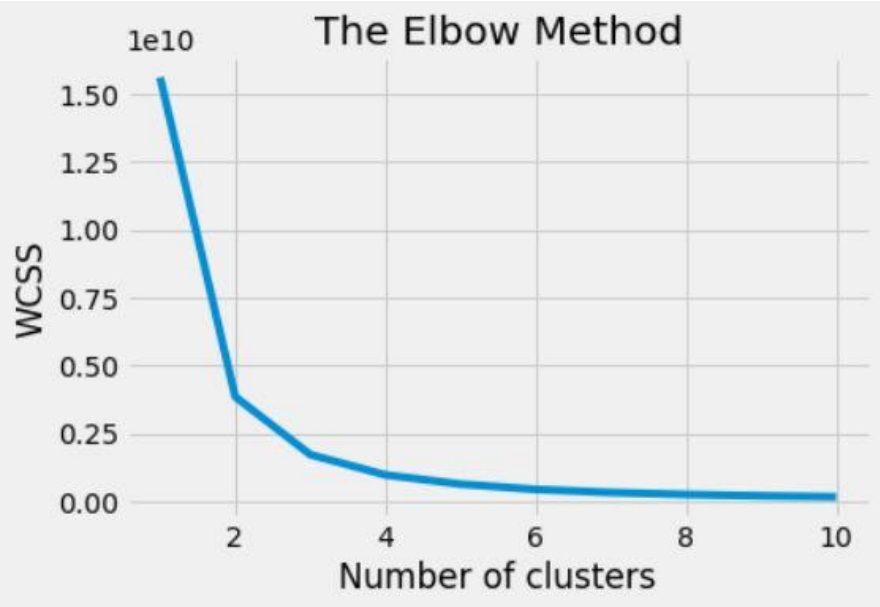



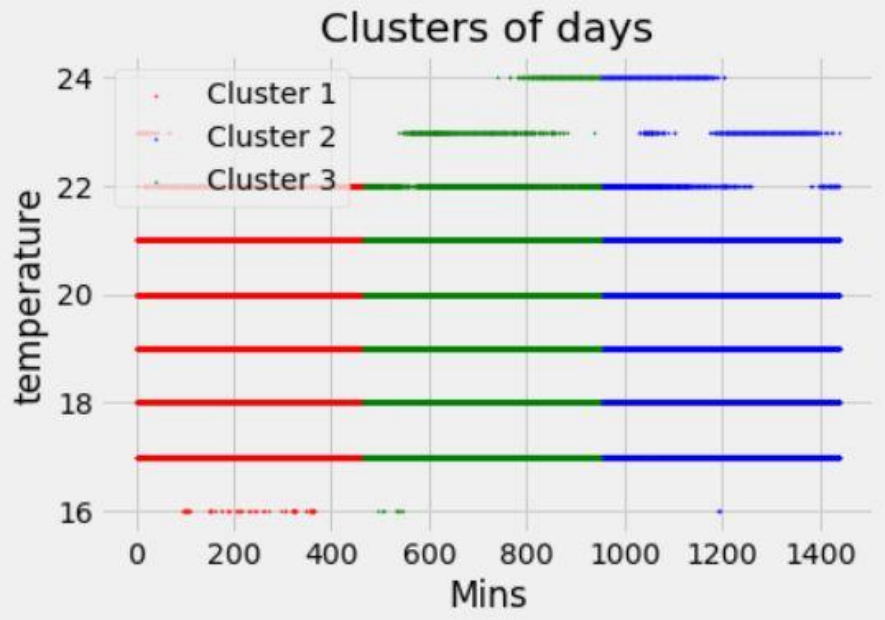
Elbow method and clusterung for **Friday**:



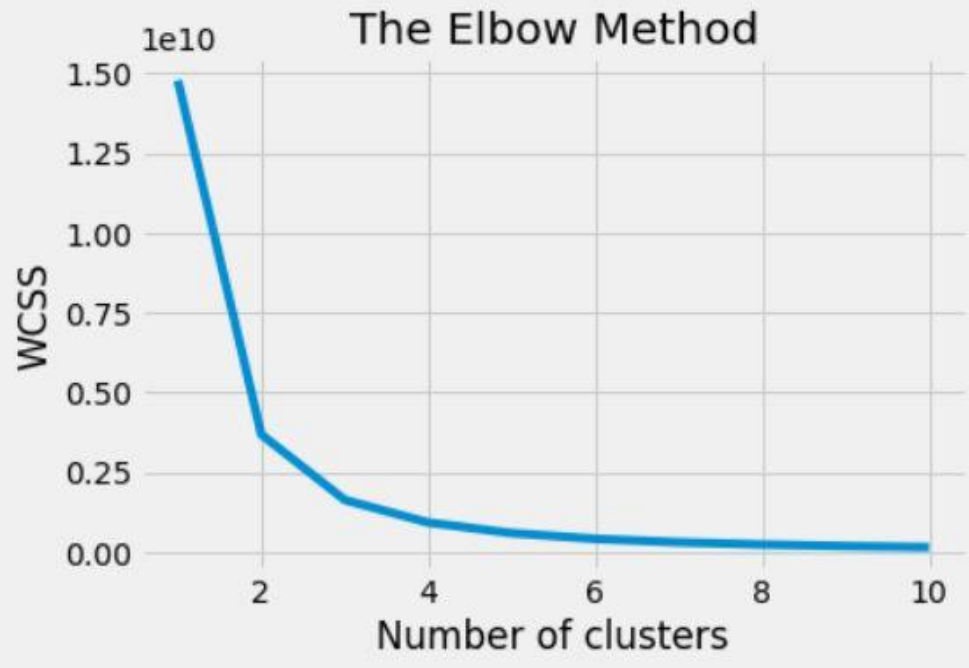



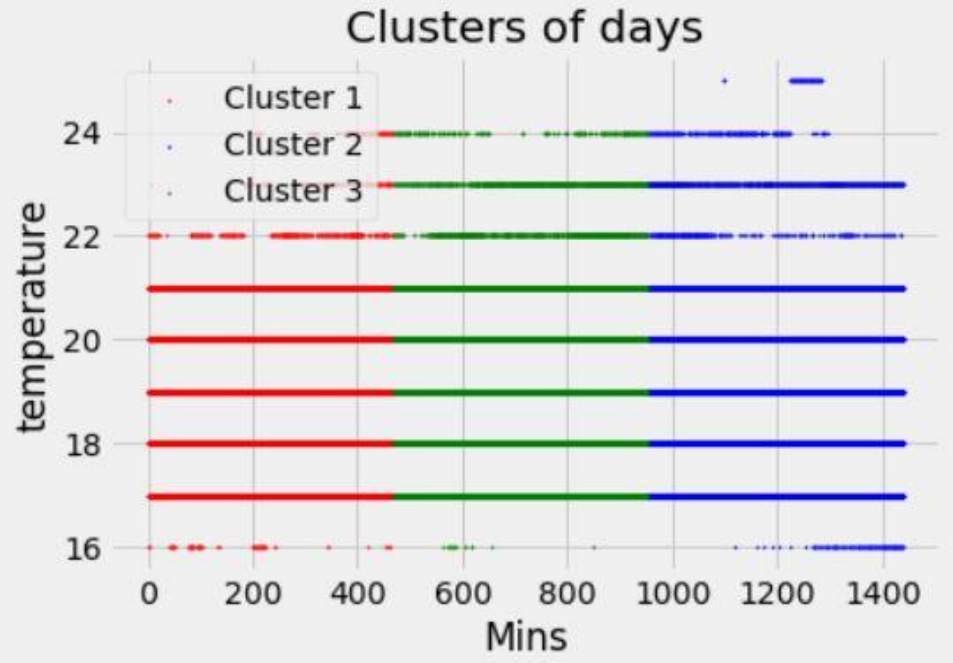
Elbow method and clusterung for **Saturday**:



Elbow method and clusterung for **Sunday**:



# CHAPTER 6

## CONCLUSION

Traditionally, weather forecasting (Temperature and Humidity) has always been performed by physically simulating the atmosphere as a fluid and then the current state of the atmosphere would be sampled. In the previous system the future state of the atmosphere is computed by solving numerical equations of thermodynamics. But this model is sometimes unstable under disturbances and uncertainties while measuring the initial conditions of the atmosphere. This leads to an incomplete understanding of the atmospheric processes, so it restricts weather prediction.

Our proposed solution of using Machine learning for weather predicting is relatively robust to most atmospheric disturbances when compared to traditional methods. Another advantage of using machine learning is that it is not dependent on the physical laws of atmospheric processes. In the long run weather prediction using Machine Learning has a lot of advantages and thus it should be used globally.

# CHAPTER 7

## FUTURE ENHANCEMENT

Weather forecasts are becoming more detailed, more accurate and are providing the information needed to make sound decisions to protect life and property. Technological advances, such as apps, are making weather information more accessible and immediately alerting those in harm's way.

The current version of Weather Prediction that we have developed is still premature. This implies that there are still many limitations that can be resolved and improved. One of the biggest limitations right now is, that the location has to be chosen from the list of places the application is bound to. This can be improved if we use web scraping tools to automatically get the weather data, for various locations, from the internet and then input it into the database. Another enhancement that can be done is the automatic validation of longitude and latitude coordinates. Another improvement that can be done is to beautify the UI to make it more appealing to the younger generation. Future enhancements will make our Weather Prediction more flexible, user friendly and thus it will be more appealing to a wide range of audience.

# REFERENCES

1. A. Payne and P. Frow, "A strategic framework for customer relationship management," *J. Marketing*, vol. 69, no. 4, pp. 167–176, Oct. 2005.

2. L. D. Xu, "Enterprise systems: State-of-the-art and future trends," *IEEE Trans. Ind. Inf.*, vol. 7, no. 4, pp. 630–640, Nov. 2011.

3. A. Berson, K. Thearling, and S. Smith, *Building Data Mining Applications for CRM*. New York: McGraw-Hill, 1999.

4. K. Coussement and D. V. Poel, "Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques," *Expert Syst. Appl.*, vol. 34, no. 1, pp. 313–327, Jan. 2008.

5. W. Verbeke, K. Dejaeger, D. Martens, J. Hur, and B. Baesens, "New insights into churn prediction in the telecommunication sector: A profit driven data mining approach," *Eur. J. Oper. Res.*, vol. 218, no. 1, pp. 211–229, Apr. 2012.

6. W. J. Reinartz and V. Kumar, "The impact of customer relationship characteristics on profitable lifetime duration," *J. Marketing*, vol. 67, no. 1, pp. 77–99, Jan. 2003.

7. P. Datta, B. Masand, D. R. Mani, and B. Li, "Automated cellular modeling and prediction on a large scale," *Artif. Intell. Rev.*, vol. 14, no. 6, pp. 485–502, Dec. 2000.

8. D. Popović and B.D. Bašić, "Churn prediction model in retail banking using fuzzy C-means algorithm," *Informatica*, vol. 33, no. 2, pp. 235–239, May 2009.

9.C.-P. Wei and I. T. Chiu, "Turning telecommunications call details to churn prediction: A data mining approach," *Expert Syst. Appl.*, vol. 23, no. 2, pp. 103–112, Aug. 2002.

10.M. Owczarczuk, "Churn models for prepaid customers in the cellular telecommunication industry using large datamarts," *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4710–4712, Jun. 2010.

11.Weather Forecasting Using Deep Learning Techniques - IEEE Conference Publication, ieeexplore.ieee.org/document/7415154.

12."Weather." Kaggle, www.kaggle.com/tags/weather.

13."Weather Forecasting." Kaggle, www.kaggle.com/questions-and-answers/27537.