



Department Of Computer Science & Engineering

SOUTHEAST UNIVERSITY

CSE4000: Research Methodology

Weather Forecasting Using Machine Learning Algorithms

A dissertation submitted to the Southeast University in partial fulfillment of the requirement for the degree of B. Sc. in Computer Science & Engineering

Submitted By

Abul Khair

ID: 2017100000008

Jobaer Bhuiyan

ID: 2017000000144

Ibrahim

ID: 2017000000117

Supervised By

Ms. Maksuda Rabeya

Lecturer, Department of CSE

Southeast University

Copyright© Year 2021

October 2021

Letter of Transmittal

October 21, 2021

The Chairman,

Department of Computer Science & Engineering,

Southeast University,

Banani, Dhaka.

Through: Supervisor, Ms. Maksuda Rabeya

Subject: Submission of Research Methodology (CSE4000) report.

Dear Madam,

It is certainly a remarkable satisfaction that we are able to present the outcome of our research on **Weather Forecasting Using Machine Learning Algorithms**. All of us worked amicably to put together this file. All the data on this file is particularly based at the data collected from the web through the internet, associated journal, and simple principle of AI concepts, numerous kinds of data & algorithms. This file is the result of the information which has been received from the respective course. We enthusiastically desire that you'll discover this file really well worth reading. If there are any queries or problems which you would really like to explain, kindly teach us for the betterment. Hope you'll respect our efforts and pardon our insignificant errors. Thank you on your co-operation.

Sincerely yours

Abul Khair

ID: 2017100000008

Batch: 46th

Program: BSc in CSE

Jobaer Bhuiyan

ID: 2017000000144

Batch: 45th

Program: BSc in CSE

Ibrahim

ID: 2017000000117

Batch: 45th

Program: BSc in CSE

Supervisor

Ms. Maksuda Rabeya

Lecturer, Department of CSE

Southeast University

Certificate

This is to certify that the research report on “**Weather Forecasting Using Machine Learning Algorithms**” has been submitted to the respected member of the board of examiner of the School of Science and Engineering in partial success of the necessities for the degree of Bachelor of Science in Computer Science & Engineering by the following students and has been typical as satisfactory.

This Paper has been performed below my guidance.

Author

Abul Khair

ID: 2017100000008

Batch: 46th

Program: BSc in CSE

Jobaer Bhuiyan

ID: 2017000000144

Batch: 45th

Program: BSc in CSE

Ibrahim

ID: 2017000000117

Batch: 45th

Program: BSc in CSE

Approved by the Supervisor

MS. Maksuda Rabeya

Lecturer, Department of CSE

Southeast University

Abstract

The weather has an effect on nearly each perspective of our day by day lives. Life would be much simpler in case we might control the weather. Until then, we'll need to settle for attempting to foresee weather but weather forecast is exceptionally unusual as indeed a small alter within the surface and air properties can intensely affect the weather. Common weather estimates, as we all know, are not all that exact as they attempt to foresee the weather conditions of expansive regions for a huge period of time as the apparatuses or mediums utilized to foresee these weather conditions are not exact sufficient. They utilize meteorological and climate information from huge regions and integrate this information into different machine learning algorithms.

Hence these weather estimates fall flat to be exact for littler areas of an expansive city. As a results day by day weather estimates we get from portable applications or broadcasts are based on bigger regions that may be less exact for a specific region of a city. To solve the less exact weather forecast issue, this research proposal centers on developing a show for precipitation forecasting with the parameters such as Temperature, Wind Speed, Wind Course, Sea level, and Humidity which are the variables that affect the result at the specific spot of intrigued.

This consider point to display a research proposal that combines hyper-accurate estimates, counting hour-by-hour precipitation forecast with customs able data to the road level utilizing supervised machine learning algorithms, Long Short Terms Memory (LSTM), and Linear Regression (LR) and bolstering authentic weather information from the past 40 a long time. The execution of these algorithms to evaluated by comparing their results with each other to find the leading algorithm suited for this research. The test results appear that the Repetitive Neural Network (RNN) models excel the linear regression show in precision and demonstrate that RNN models can be an effective way for weather forecasting.

Acknowledgments

At first and above all, we need to renowned almighty ALLAH who presenting us this possibility and granting us the functionality to finish this research paintings and ultimately completed researching a file with none problem so far.

We are pleased to renowned supervisor **Ms. Maksuda Rabeya, Lecturer, Department of Computer Science & Engineering.** Whose heat support, insightful course and effective concept in the course of this exam paintings relatively helped us to complete an exploration effectively.

We would also like to thank all the faculty members at Southeast University Department of CSE for their steering and guide. We could also want to thank all of my friends for his or her wonderful suggestions, encouragement, and feedback.

Last but not least, we would really like to specific our gratitude to our parents, who've usually supplied us with intellectual guide and encouragement. Finally, we're thankful to have this kind of first-rate institution member for his or her interesting conversations, sleepless nights collectively to finish this assignment on time, and all of the times we've had over the previous few months. We can whole this document on time thanks to all your assist and encouragement.

Table of Contents

Abstract.....	III
Acknowledgments.....	IV
List of Figures.....	VI
Chapter 1- Introduction.....	7
1.1 Problem Statement.....	8
1.2 Objective.....	9
1.3 Motivation.....	9
 Chapter 2- Literature Review.....	 10-11
2.1 Limitation of Weather forecasting.....	12
Chapter 3- Research Methodology.....	13
3.1 Data Collection.....	14
3.2 Data Preprocessing.....	14
3.3 Feature Selection.....	14
3.4 Train Test Split.....	14
3.5 Supervised Learning Algorithm.....	15-16
3.6 Linear Regression.....	17-18
 Chapter 4- Implementation and Result.....	 19-26
Chapter 5- Conclusions.....	27
 References.....	 28

Table of Figures

Figure 1: Flow chart of research methodology.....	13
Figure 2: Different hyperplane of SVM.....	16
Figure 3: Non-linear Support vector Machine (SVM).....	16
Figure 4: Linear Regression in Machine Learning.....	18
Figure 5: Weather forecasting Dataset.....	20
Figure 6: Import Libraries & Data.....	21
Figure 7: Graph our Dataset.....	22
Figure 8: Define X Values	23
Figure 9: Model the training set & predict the test set results...	24
Figure 10: Evaluate the model & plot the result.....	25
Figure 11: Predicted values output & Difference.....	26

Chapter 1

Introduction

Weather forecasting is the application of logical methods and innovation to foresee the conditions of the environment at a certain area and time. Weather Forecasting in ancient time is carried out by hand, utilizing changes in barometric weight, current weather conditions, and sky condition or cloud cover, weather estimating presently depends on computer-based models that take numerous air components in to accounting presently depends on computer-based models that take numerous air variables into consideration.

For a long time, the researcher had attempted to set up a linear relationship between the input weather information properties and the comparing target property. But the revelation of nonlinearity inside distinctive properties of weather information, the center has moved towards the nonlinear forecast of the weather. Weather forecasts are made by collecting quantitative information around the current state and past slant of the environment and utilizing logical understanding of climatic processes to foresee how the environment will advance. The weather warning is imperative for the security of life and property. Rain forecasts can be utilized by farmers.

In arrange to analyze how the distinctive machine learning methods will perform within the estimating of weather. We have prepared distinctive sorts of machine learning models on information collected from the air terminal weather station of a few cities.

1.1 Problem Statement

Heavy precipitation can lead to various dangers, for occurrence: flooding, including danger to human life, harm to structures and system, and misfortune of products and domesticated animals. Which can compromise human life, disturbed transport and trading, and cause harm to structures and establishment. Where overpowering precipitation happens with high breezes, risk to ranger benefit crops is high.

In the case of starting treatment of patients, the probability of survival has expanded essentially with early determination of cancer. With appropriate tumor classification, pointless treatment can be maintained a strategic distance from. Each volume should be treated differently. Hence, in case there's no proper diagnosis at that point there's a high risk of death for the persistent. Adjust determination of cancer and classification of tumors in kind and dangerous tumors is a range of investigation.

For example, in case we consider a range influenced by tropical cyclone the basic impacts of tropical cyclone incorporate heavy rain, strong wind, tremendous storm floods near landfall, and tornadoes. The destruction from a tropical cyclone, for case, a sea tempest or storm, depends for the foremost portion on its control, its measure, and its range. Tropical tornados act to empty woods shade and moreover change the scene near beach front zones, by moving and reshaping sand edges and causing wide disintegration along the float. In fact, indeed well inland, overpowering precipitation can prompt mudslides and torrential slides in tough districts. Their assets can be identified after a few times by concentrate the merging of the Oxygen isotope inside caverns interior the region of typhoons' ways. So, we are giving distant better much better higher stronger an improved a distant better way to induce precise forecasts. As specified above, the benefits of identifying critical features of machine learning, complex information sets, play a critical part in forecasting of weather. Since the most excellent results can be accomplished with designing learning algorithms, we should utilize these techniques to mindful individuals from normal disasters. Usually since learning designing algorithms can give more precise results. Separated from this, the results are accomplished at a brief time and individuals get sufficient time to do arrangements or to escape from that place.

1.2 Objective

Our research aims to foresee the Weather and Environment conditions utilizing the previous dataset of the weather forecasting with a center on progressing the precision of forecast. This will increment the precision of the weather forecast and we'll get exact comes about than the conventional methods. Our dataset comprises of max and min. temperature of ordinary from the particular area. When gathering datasets to allow to the models there are sure parameters which are called as requested data which incorporates: snow, rainstorm, rain, mist, cloudy, for the foremost portion cloudy, halfway shady, scattered mists, and clear. In this way our point is to provide accurate result in arrange to provide correct forecast of weather for future so in basic conditions individuals can be mindful of upcoming common calamities.

1.3 Motivation

- Weather forecasting community needs to supply valuable weather forecasts and communicate estimate data effectively.
- Need to understand clients estimate data needs, recognitions, interpretations, preferences and uses.
- Audience for National Weather Service estimates incorporates mediators, particular user groups and public.
- Weather forecasts are inherently dubious and numerous clients realize estimates are defective but most current weather estimate data given to the open is deterministic.
- There is an interest in giving instability data but it is challenging to do so effectively.
- Weather forecasters see at current state of the weather and estimate maps and include their individual encounter to come up with an estimate and to issue warnings.

Chapter 2

Literature Review

Weather forecasting has been one of the foremost challenging troubles around the world since of both its practical value in prevalent scope for logical study and meteorology. Weather could be a persistent, energetic, multidimensional chaotic prepare, and data-intensive and these properties make weather forecasting a fortifying challenge. It is one of the foremost imperious and requesting operational duties that must be carried out by numerous meteorological services all over the globe. Different organizations / specialists in India and abroad have done illustrating utilizing upheld time arrangement information control. The different technique measurement decomposition models, Exponential smoothing models, numerous trainings have taken put inside the investigation of design and circulation of precipitation in numerous locales of the world. Completely changed time arrangement methods with distinctive purposes are utilized to explore weather data in numerous distinctive literatures. Exact and opportune weather forecasting may be a major challenge for the logical investigate. Weather forecast displaying includes a combination of numerous computer models, perceptions and associate of patterns and plans. Utilizing these methods, for all intents and purposes accurate forecasts can be made up. Relapse could be a statistical test procedure and it must be broadly utilized in numerous commerce, behavioral sciences, social and climate recasting and numerous other regions.

Consistency of weather figures, consistency of weather with a numerical arrangement of measurements that control development and weather alter. Numerous weathers determining procedures, in expansion to being utilized for short term weather forecasting, are being utilized in research considers such as discuss contamination and the impacts of greenhouse gases on worldwide weather change.

The first dynamic weather forecasting show comprised of as it were one layer and could subsequently appear as it were brief varieties within the vertical structure of the environment. Computers presently allow the development of multilevel models that can solve vertical varieties, temperature and humidity. These multi-mile species foresee the fundamental weather variances of expansive development scales.

They used a logistic regression approach, in which just two values are utilized: 0 for no precipitation and 1 for precipitation. It implies that maximum and minimum temperatures, as well as evening humidity, have a positive effect on precipitation, but dissipation and morning humidity includes a negative effect. Maximum temperature, agreeing to meteorologists, ought to have a negative effect on precipitation. The research's deviation measurement infers that the information fit the show effectively. Indeed however, at the 10percent levels of significance, appear measurement shows that the fit is destitute. At that point they assess for exceptions within the information utilizing the Generalized Standardized Pearson Residuals demonstrative. After that, they expel the exceptions from the information and re-fit the logistic regression show by utilizing the remaining information. The changed result shows that minimum temperatures and evening humidity have a positive effect on precipitation, though maximum temperatures, dissipation, and morning humidity have a negative effect. They moreover feel that disposing of the distant things brought about in a significant change in show fitting as appeared by appear measurement. At last, they explore the validity of their outfitted show utilizing cross validation investigation.

In addition, presently and once more it is seen that wrong weather forecast is making various issues. Current satellites, radar, sensors, combined with rapidly decreasing costs of putting absent and circulating information have catalyzed an impact in amounts of weather data available for considers. Most work in weather to date depends on the utilization of generative techniques, where the weather frame works are reproduced by implies of numerical procedures. Scarcely any inquires around have been displayed within the field of current weather estimate. In show disdain toward of the reality that it isn't completely depend upon computerized thinking system or machine learning based. They are furthermore doing work based on few data set. By utilizing these data exact weather temperature is past the domain of creative ability to expect to expect. An expansive parcel investigates do not seek after machine learning process. In spite of the reality that in case a few examinations utilize machine learning system however their procedure, calculation isn't refreshed. Also, as of presently there's no investigation has been done on temperature estimate of Bangladesh. In this paper we have proposed a method of machine figuring out how to predict the temperature of all over Bangladesh. This will push everyone to express temperature all over all through the year.

2.1 Limitation of Weather forecasting

- The needs and importance of weather forecasting is apparent, in any case we are still limited in what we will foresee. We have minimal control over weather and ultimately the environment is as well unusual to ever know with 100% precision what is progressing to happen.
- Weather forecasting too is valuable for different purposes- aviation, shipping, fisheries, flood forecasting, cyclone warning and numerous other uncommon users other than forecasts for the general public.
- Weather forecasting can help us avoid misfortune of life, property damage and indeed makes a difference us prevent weather related sicknesses such a hyperthermia.
- Weather forecasting is very critical to makes a difference individuals get ready for an assortment of weather conditions, a few of which can be unsafe to our health, others that can cause unsafe driving conditions, and a few that can cause major property damage.
- A forecast makes a difference us to form choices day by day based on weather conditions. It can be as basic as what clothing to wear or as complicated as stocking up on food and supplies to get ready for extreme weather.

Chapter 3

Research Methodology

The following steps are used to achieve the objectives of this paper—

- Setup
- Data collection
- Data preprocessing and
- Training models

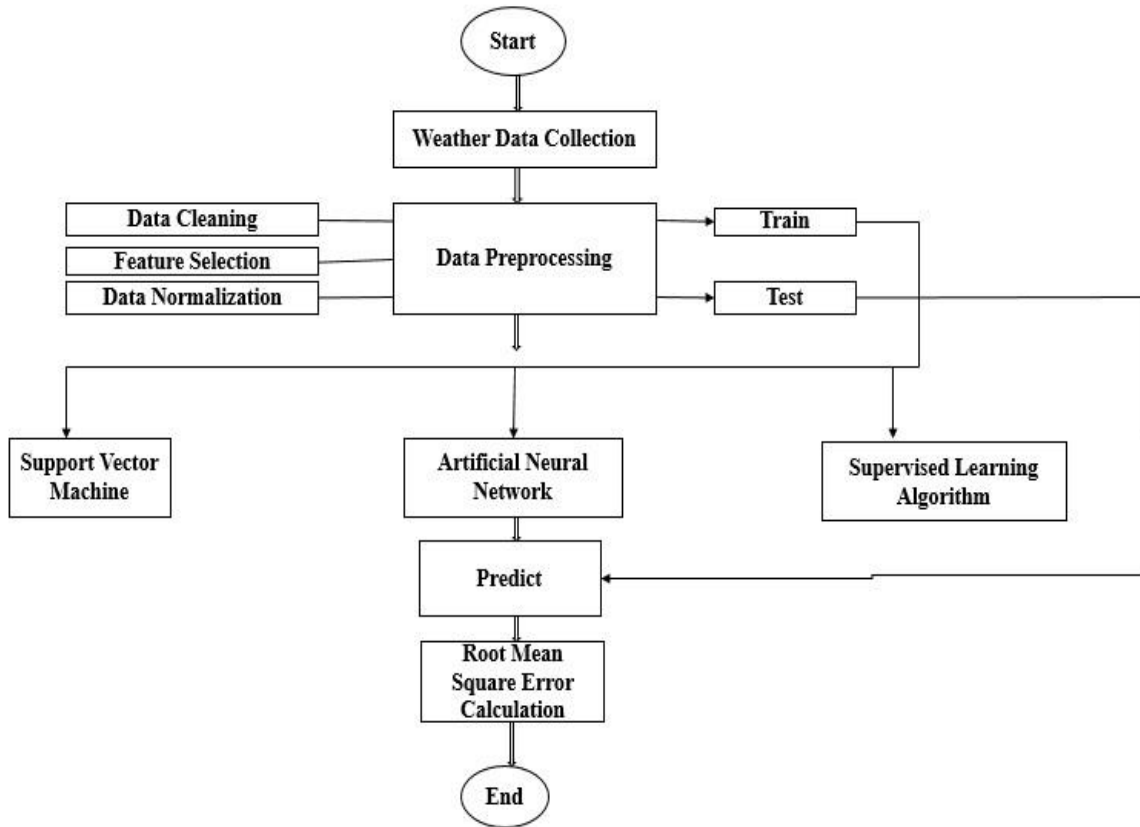


Figure 1: Flow chart of research methodology

3.1 Data Collection

In order to analyze different machine learning procedures. We have collected information from different Bangladesh weather. We collected a level www.visualcrossing.com that contained news articles on Weather Forecasting from.

3.2 Data Preprocessing

The next step is preprocessing of information. In this step, information is stacked into appropriate put & prepared to utilize in preparing of machine learning. Information is separated into 2 parts. The first part of the information is training information and other part of the information is testing information. These information sets are utilized for progressing model's performance.

3.3 Feature Selection

The relationship between different columns of the weather information is considered and their framework is produced appearing linear dependence between diverse columns. Columns with high linear dependence don't give critical data in forecast hence expelled from the dataset.

3.4 Train Test Split

In machine learning, Train test split could be a work of the scalar model inclination to separate information arrays into two subdivisions: training data and test data. With this work, there's no have to be isolated the dataset physically. However, the scalar train test split will get random division for the two subsets. Sometime recently giving information in machine learning, we need to split the dataset for train and test form. We have tested and assessed the linear regression model on the precipitation dataset and the precision of the model is 0.066% further we'll apply the multiple linear regression show to discover the precision of the show on the dataset.

3.5 Supervised Learning Algorithm

This segment has been highlighted the SVM algorithm since it works well for this dataset. The support vector machine (SVM) algorithm can be utilized for both classification and regression issues. In any case, SVM is very prevalent for generally complex types of little or medium classification datasets. In this algorithm, the information points are isolated by a hyperplane, in which case the part chooses the hyperplane's perceptions. A few basic components can be found for the SVM, such as Support Vector, Hyperplane, and Margin. Support vectors are the information points at the edges of distinctive classes of datasets that separate between classes. The hyperplane may be a decision-making line that separates between multiple levels of information. Margin is the separate between multiple levels of information. The separate between the support vectors decides the margin. The higher the margin between the two classes, the better the margin.

On the other hand, there are distinctive sorts of support vector machine parts, for occurrence, linear part, Polynomial part, and Radial Basis Function Part. In case we plot more than one variable in a typical scatter plot, in numerous cases, that plot cannot isolated two or more information classes. The part of the support vector machine may be a special sort of approach that can change over lower-dimensional information into higher dimensional space and separate between classes. In any case, the SVM as a rule works in another way for the linear problem. To begin with, multiple hyperplanes are produced within the to begin with picture in Figure.2. (Left side of Figure.2) ready to see that three hyperplanes have been produced for classification. At that point the classification error of each hyperplane is calculated, it is found which hyperplane has the most reduced error which hyperplane is chosen for classification.

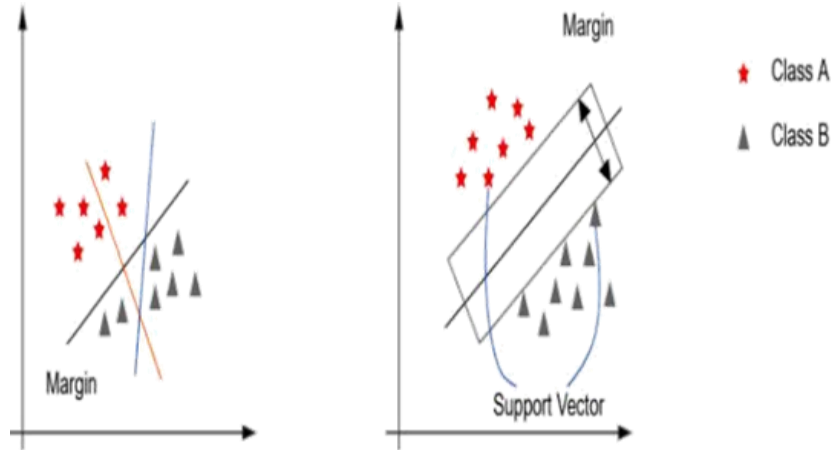


Figure 2: Different hyperplane of SVM

Many information cannot be found to be linearly isolated for classification, in which case the non-linear part changes the information into higher dimensions, making it valuable for classification. Figure.3 appears the non-linear support vector machine. The point to arrangement and results are shown in Figure.3.

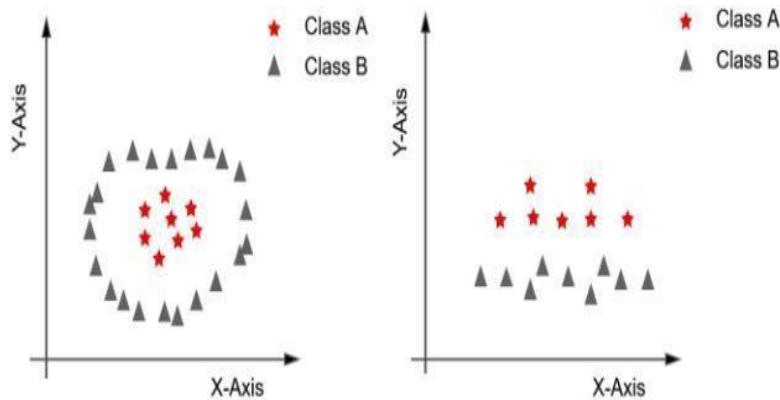


Figure 3: Non-linear Support vector Machine (SVM)

3.6 Linear Regression

Linear regression investigation is utilized to foresee the value of a variable based on the value of another variable. The variable you need to foresee is called the dependent variable. The variable you're utilizing to foresee the other variable's value is called the independent variable.

This form of investigation gauges the coefficients of the linear equation, including one or more independent variables that best foresee the value of the dependent variable. Linear regression fits a straight line or surface that minimizes the inconsistencies between predicted and real yield values.

There are basic linear regression calculators that utilize a “least squares” strategy to find the best-fit line for a set of combined information. You at that point gauge the esteem of X (dependent variable) from Y (independent variable).

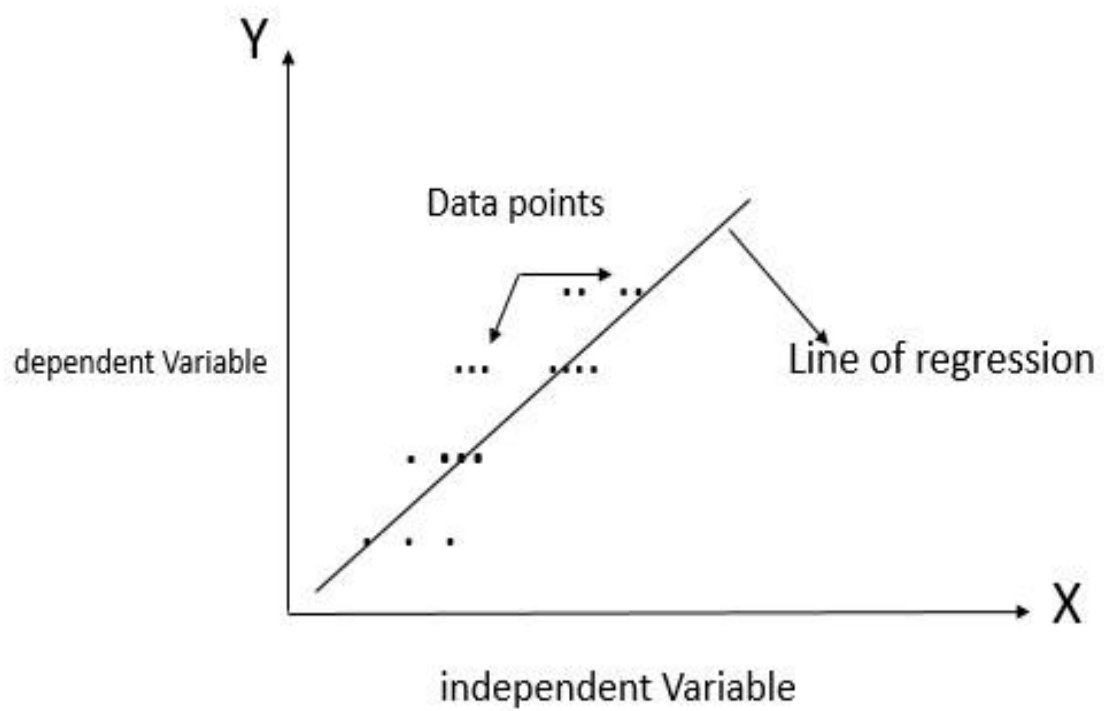


Figure 4: Linear Regression in Machine Learning

Chapter 4

Implementation & Result

Implementation

Now we are going to implement our research situation utilizing the Python programming language. We will utilize Python and the execution will be finished utilizing Jupyter, which could be a result of Google Investigate. Jupyter may be an organize that grant anyone to utilize all Python libraries in their code with no obstacle. Many stages are taken after to finish the investigation. We have attempted to implement that to prove the realness of our research scenario.

Dataset

We collect a dataset from Weather Data Online. (2021). Bangladesh. Recovered from <https://www.visualcrossing.com/weather/weather-data-services#/viewData>. We convert our dataset into two sections, one is the Train dataset and another one is for testing. Here we'll implement our Linear Regression on our train dataset and will utilize a test dataset for forecasting the precision.

Our Dataset

	A	B	C	D	E	F	G	H	I	J	K	L	M
	Date time	Maximum Temperature	Minimum Temperature	Temperature	Heat Index	Chance Precipitation (%)	Wind Speed	Wind Direction	Wind Gust	Visibility	Cloud Cover	Relative Humidity	Conditions
1													
2	9/24/2021	93.8	83.9	89.1	102.3	0	6.4	299.8	10.3	14.9	9	61.2	Clear
3	9/25/2021	94.4	81.1	85.5	104.1	14.3	5.8	209.2	13.9	12.6	54.5	71.2	Partially cloudy
4	9/26/2021	93.8	79.2	85.1	104	57	5.5	182	17.7	4.8	59.3	71.5	Rain, Partially cloudy
5	9/27/2021	82.7	77.6	80	89.6	81	4.9	119.2	16.3	2.4	93.9	85.8	Rain, Overcast
6	9/28/2021	87.5	77.3	81.5	95.5	52.4	4.6	121.5	11.9	15	60.1	79.1	Rain, Partially cloudy
7	9/29/2021	92.9	77.6	84	100.8	47.5	6.2	178.2	14.1	15	80.5	73.6	Overcast
8	9/30/2021	93.8	78.7	84.9	102	61.9	4.4	206.5	10.6	15	75.3	72	Rain, Overcast
9	10/1/2021	90.9	78.5	84.2	101.2	52.4	6.2	218.2	18.3	4.9	79.1	75.6	Rain, Overcast
10	10/2/2021	91.1	77.6	83.1	101.5	61.9	8.8	209	19.9	2.4	82.8	78.9	Rain, Overcast
11	10/3/2021	81.2	76.9	78.6	87.3	66.6	7.5	183.4	18.8	2.9	98.4	90.6	Rain, Overcast
12	10/4/2021	81.6	74.7	77	87.3	76.1	7.5	178.8	17.4	3	83.1	90.2	Rain, Overcast
13	10/5/2021	87.3	74.9	80.2	95.3	71.4	6.7	156	9.4	5.3	40	80.6	Rain, Partially cloudy
14	10/6/2021	89	76.2	81.7	97.2	85.6	6	172.8	15.2	15	20.2	78.5	Rain
15	10/7/2021	89.5	76.6	81.5	99	76.1	7.3	175	14.8	2.3	76.7	81.4	Rain, Overcast
16	10/8/2021	88.6	77.6	82	97.7	66.6	6.6	158.5	13.2	8.7	76.3	78.5	Rain, Overcast
17	10/9/2021	90.1	77.3	82.5	98.8	76.1	6.7	161.7	12.5	5.5	85.9	77.7	Rain, Overcast
18	10/10/2021	94.2	81.9	86.5	98.8	23.7	3.9	177.8	6.3	15	49.2	59.1	Partially cloudy
19	10/11/2021	94.9	80	86.8	99.9	9.5	4.2	125.4	7.6	15	9.7	58.7	Clear

Figure 5: Weather forecasting Dataset

Importing Libraries & Data

We use .csv file. Since .csv files are a common file organize for exchanging and putting away data. The capacity to studied, manipulate, and type in data to and from .csv files utilizing Python could be a key ability to ace for any data researcher. We collected the data and imported several libraries to import our dataset. Pandas could be a machine learning instrument for data cleaning and analysis too. Our dataset has to be visualized and cleaned and Pandas may be a great instrument for this purpose.

```
In [1]: #Import Libraries
import pandas as pd
import numpy as np

In [2]: #Import dataset
data_df=pd.read_csv('D:forecast.csv')

In [3]: data_df.head()

Out[3]:
```

	Maximum Temperature	Minimum Temperature	Temperature	Heat Index	Chance Precipitation	Wind Speed	Wind Direction	Wind Gust	Visibility	Cloud Cover	Relative Humidity
0	93.8	83.9	89.1	102.3	0.0	6.4	299.8	10.3	14.9	9.0	61.2
1	94.4	81.1	85.5	104.1	14.3	5.8	209.2	13.9	12.6	54.5	71.2
2	93.8	79.2	85.1	104.0	57.0	5.5	182.0	17.7	4.8	59.3	71.5
3	82.7	77.6	80.0	89.6	81.0	4.9	119.2	16.3	2.4	93.9	85.8
4	87.5	77.3	81.5	95.5	52.4	4.6	121.5	11.9	15.0	60.1	79.1

```
In [4]: #Define x and y
x=data_df.drop(['Chance Precipitation'],axis=1).values
y=data_df['Chance Precipitation'].values

In [5]: print(x)

[[ 93.8  83.9  89.1 102.3  6.4 299.8 10.3 14.9  9.  61.2]
 [ 94.4  81.1  85.5 104.1  5.8 209.2 13.9 12.6 54.5 71.2]]
```

Figure 6: Import Libraries & Data

Graph Our Dataset

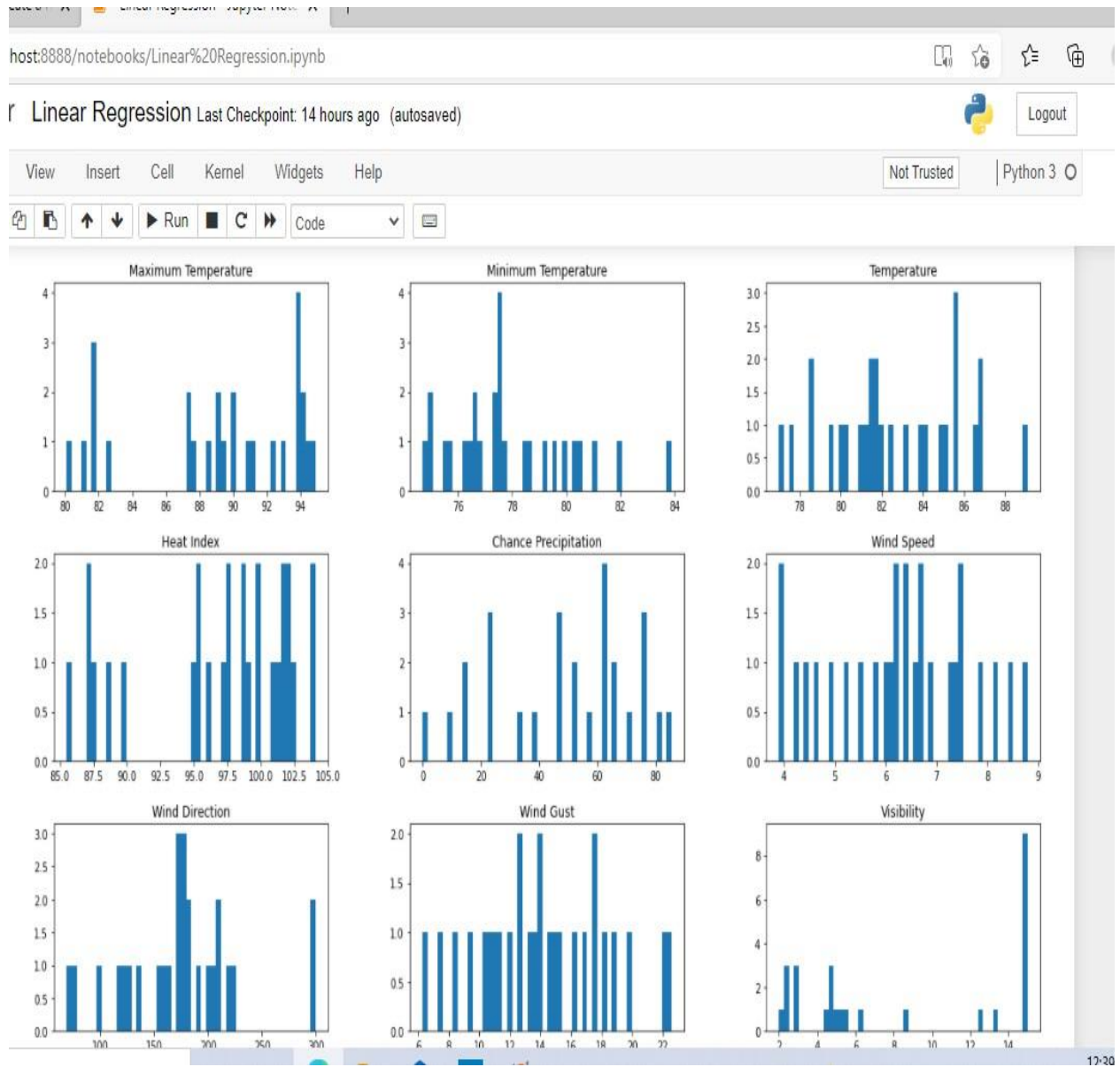
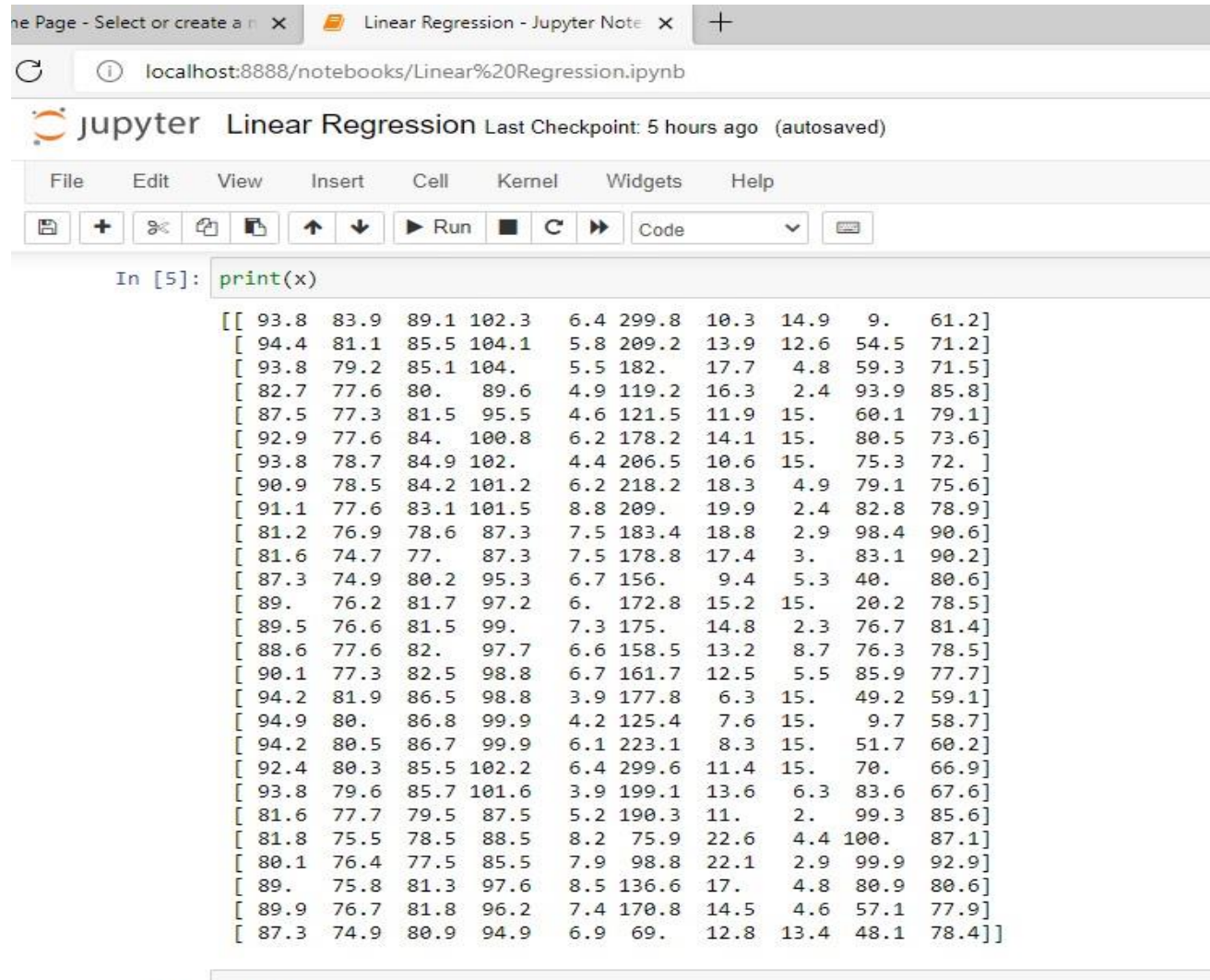


Figure 7: Graph our Dataset

Define X & Y



The screenshot shows a Jupyter Notebook window titled "Linear Regression - Jupyter Note". The address bar indicates the notebook is running on localhost:8888. The notebook interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for saving, running, and other actions. The current cell is a code cell with the following content:

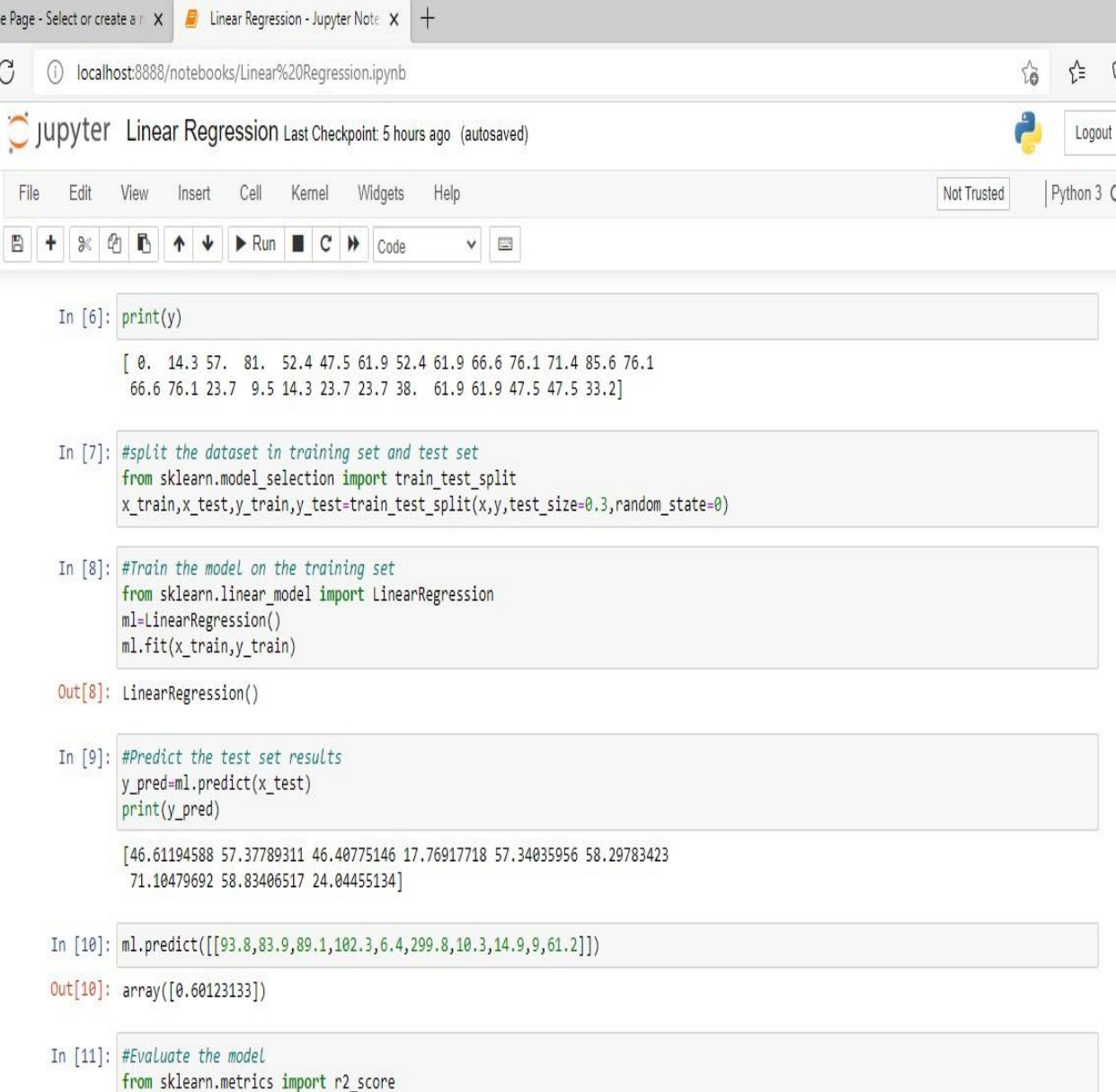
```
In [5]: print(x)
```

The output of the cell is a 2D array of numerical values, displayed as follows:

```
[[ 93.8  83.9  89.1 102.3   6.4 299.8  10.3  14.9   9.   61.2]
 [ 94.4  81.1  85.5 104.1   5.8 209.2  13.9  12.6  54.5  71.2]
 [ 93.8  79.2  85.1 104.   5.5 182.   17.7   4.8  59.3  71.5]
 [ 82.7  77.6  80.   89.6   4.9 119.2  16.3   2.4  93.9  85.8]
 [ 87.5  77.3  81.5  95.5   4.6 121.5  11.9  15.   60.1  79.1]
 [ 92.9  77.6  84.   100.8   6.2 178.2  14.1  15.   80.5  73.6]
 [ 93.8  78.7  84.9 102.   4.4 206.5  10.6  15.   75.3  72. ]
 [ 90.9  78.5  84.2 101.2   6.2 218.2  18.3   4.9  79.1  75.6]
 [ 91.1  77.6  83.1 101.5   8.8 209.   19.9   2.4  82.8  78.9]
 [ 81.2  76.9  78.6  87.3   7.5 183.4  18.8   2.9  98.4  90.6]
 [ 81.6  74.7  77.   87.3   7.5 178.8  17.4   3.   83.1  90.2]
 [ 87.3  74.9  80.2  95.3   6.7 156.   9.4   5.3  40.   80.6]
 [ 89.   76.2  81.7  97.2   6.   172.8  15.2  15.   20.2  78.5]
 [ 89.5  76.6  81.5  99.   7.3 175.   14.8   2.3  76.7  81.4]
 [ 88.6  77.6  82.   97.7   6.6 158.5  13.2   8.7  76.3  78.5]
 [ 90.1  77.3  82.5  98.8   6.7 161.7  12.5   5.5  85.9  77.7]
 [ 94.2  81.9  86.5  98.8   3.9 177.8   6.3  15.   49.2  59.1]
 [ 94.9  80.   86.8  99.9   4.2 125.4   7.6  15.   9.7  58.7]
 [ 94.2  80.5  86.7  99.9   6.1 223.1   8.3  15.   51.7  60.2]
 [ 92.4  80.3  85.5 102.2   6.4 299.6  11.4  15.   70.   66.9]
 [ 93.8  79.6  85.7 101.6   3.9 199.1  13.6   6.3  83.6  67.6]
 [ 81.6  77.7  79.5  87.5   5.2 190.3  11.   2.   99.3  85.6]
 [ 81.8  75.5  78.5  88.5   8.2  75.9  22.6   4.4 100.   87.1]
 [ 80.1  76.4  77.5  85.5   7.9  98.8  22.1   2.9  99.9  92.9]
 [ 89.   75.8  81.3  97.6   8.5 136.6  17.   4.8  80.9  80.6]
 [ 89.9  76.7  81.8  96.2   7.4 170.8  14.5   4.6  57.1  77.9]
 [ 87.3  74.9  80.9  94.9   6.9  69.   12.8  13.4  48.1  78.4]]
```

Figure 8: Define X Values

Model the training set & predict the test set results



The screenshot shows a Jupyter Notebook titled "Linear Regression" running on a local host. The notebook contains several code cells for training and testing a linear regression model. The output of the first cell shows the target variable 'y'. The second cell splits the data into training and testing sets. The third cell trains the model. The fourth cell predicts the test set results. The fifth cell shows the prediction for a specific input. The sixth cell evaluates the model using the R-squared score.

```
In [6]: print(y)

[ 0. 14.3 57. 81. 52.4 47.5 61.9 52.4 61.9 66.6 76.1 71.4 85.6 76.1
 66.6 76.1 23.7 9.5 14.3 23.7 23.7 38. 61.9 61.9 47.5 47.5 33.2]

In [7]: #split the dataset in training set and test set
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=0)

In [8]: #Train the model on the training set
from sklearn.linear_model import LinearRegression
ml=LinearRegression()
ml.fit(x_train,y_train)

Out[8]: LinearRegression()

In [9]: #Predict the test set results
y_pred=ml.predict(x_test)
print(y_pred)

[46.61194588 57.37789311 46.40775146 17.76917718 57.34035956 58.29783423
 71.10479692 58.83406517 24.04455134]

In [10]: ml.predict([[93.8,83.9,89.1,102.3,6.4,299.8,10.3,14.9,9,61.2]])

Out[10]: array([0.60123133])

In [11]: #Evaluate the model
from sklearn.metrics import r2_score
```

Figure 9: Model the training set & predict the test set results

Evaluate the model & plot the result

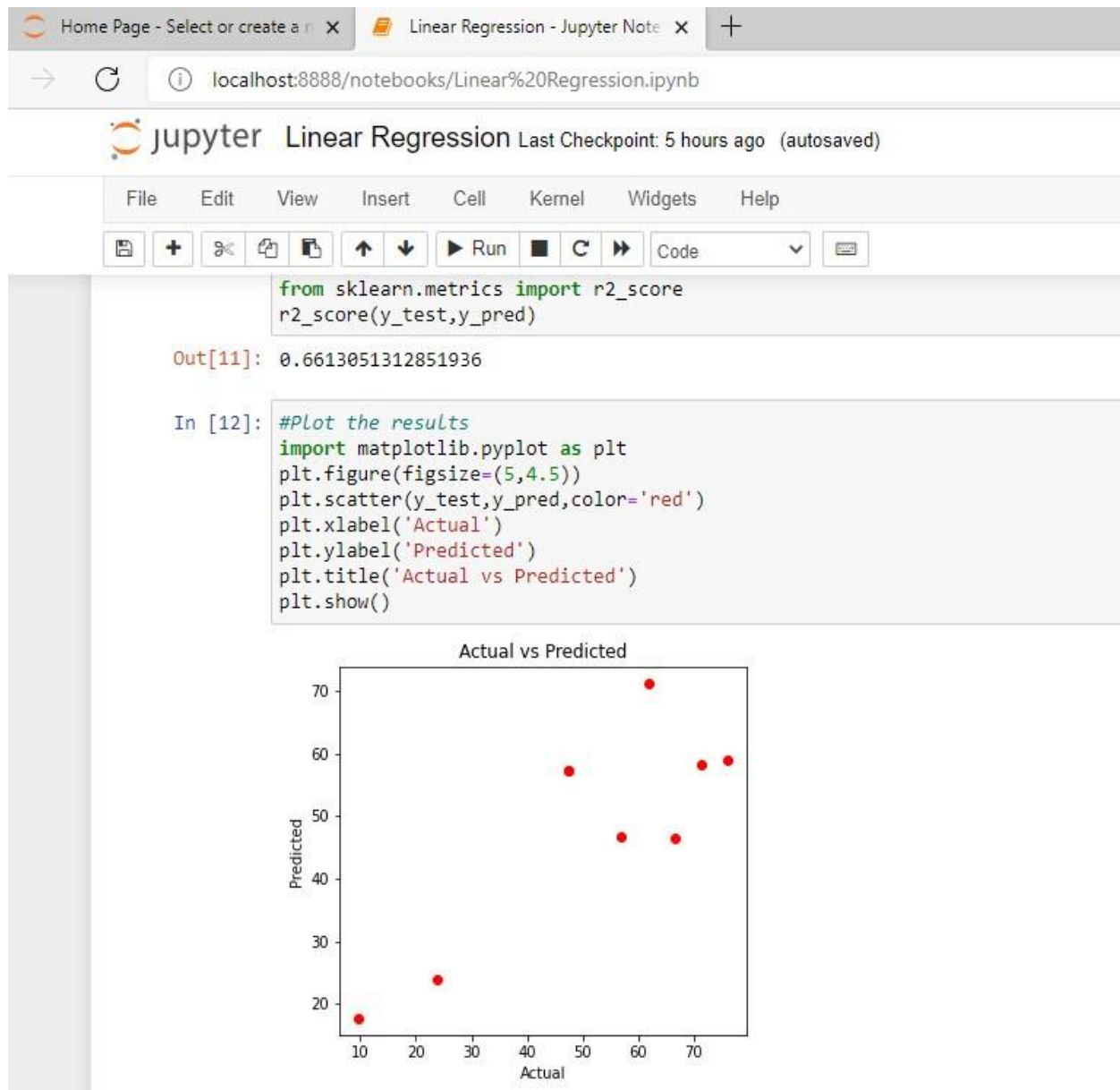


Figure 10: Evaluate the model & plot the result

Predicted values output & Difference

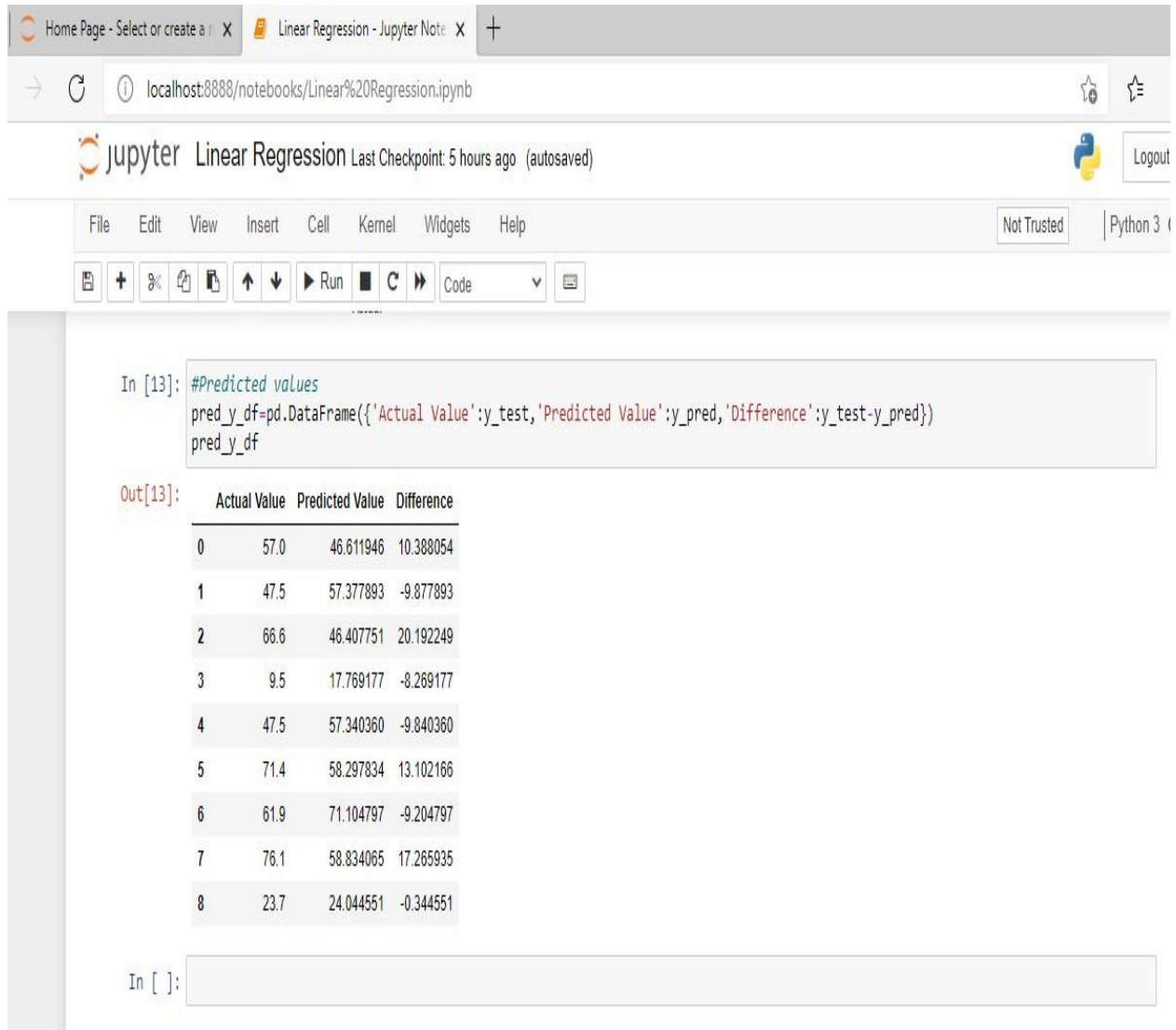


Figure 11: Predicted values output & Difference

Chapter 5

Conclusions

In this area of global warming, research in weather measurement, monitoring and forecasting are becoming more and more relevant.

Weather Forecasting plays an important role in human life, so the collection of information about the temporal dynamics of weather changes is very paramount.

This research demonstrates the design and Implementation of an affordable mini weather monitoring system that ensures portability, scalability and user friend operations which can provide data of some weather variables including temperature, humidity, wind speed, wind direction with real time data. It was calculated using the linear regression formula.

References

1. Weather Data Online. (2021). Bangladesh. Retrieved from <https://www.visualcrossing.com/weather/weather-dataservices#/viewData>
2. Beni, S. &Arora, J. (2012), —Classification and feature selection techniques in data mining||, International Journal of Engineering Research & Technology (IJERT), 1(6)
3. Kumar, R., & Khatri, R. (2016). A Weather Forecasting Model using the Data Mining Technique. International Journal of Computer Applications, 139(14).
4. A Quick Review of Machine Learning Algorithms, 2019, International Conference on Machine Learning, Big Data, Cloud and Parallel Computing.
5. https://www.ripublication.com/ijaer19/ijaerv14n4_12.pdf
6. D. B. Shank,G. Hoogenboom,R. W. McClendon "Dewpoint Temperature Prediction Using Artificial Neural Networks" Artificial Intelligence Center, University of Georgia, Athens, Georgia,JAMC,2008.
7. Meera Narvekar, Priyanca Fargose, “Daily weather forecasting using Artificial Neural Network”, International Journal of Computer Application, 2015.
8. Qing, X., Niu, Y. “Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM”. Energy, Elsevier, vol. 148(C).
9. Mark Holmstrom, Dylan Liu, Christopher Vo, “Machine Learning applied to weather forecasting”.