Machine Learning Approach to Predict Rainfall Amount of Dhaka

Submitted by

Abdul Jabbar Gazi

ID: 171-35-2062

Department of Software Engineering

Daffodil International University

Supervised by

Asif Khan Shakir

Sineor Lecturer

Department of Software Engineering

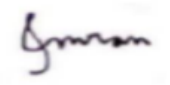
Daffodil International University

This Project report has been submitted in fulfillment of the requirements for the Degree of Bachelor of Science in Software Engineering.

APPROVAL

This thesis titled "Machine Learning Approach to Predict Rainfall Amount of Dhaka", submitted by Abdul Jabbar Gazi, ID: 171-35-2062 to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Software Engineering (SWE) and approved as to its style and contents

BOARD OF EXAMINERS



Dr. Imran Mahmud
Associate Professor and Head
Department of Software Engineering
Daffodil International University

Chairman

Md Anwar Hossen
Assistant Professor
Department of Software Engineering
Daffodil International University

Internal Examiner 1

Asif Khan Shakir
Senior Lecturer
Department of Software Engineering
Daffodil International University

Internal Examiner 2

Professor Dr M Shamim Kaiser
Institute of Information Technology
Jahangirnagar University

External Examiner

DECLARATION

I hereby state that I have taken this thesis under the supervision of Asif Khan Shakir, Senior Lecturer, Department of Software Engineering, Daffodil International University. I also acknowledge that neither this thesis nor any part of this has been submitted elsewhere for the award of any degree previously by others.

_____

Abdul Jabbar Gazi
ID: 171-35-2062
Batch 22th
Department of Software Engineering
Faculty of Science & Information Technology
Daffodil International University

Certified by:

_____

Asif Khan Shakir
Senior Lecturer
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

ACKNOWLEDGEMENT

This thesis I am representing was only possible to complete with guidance from some conscientious people. I want to thank each of them. Especially obligated to Daffodil International University for the direction and constant supervision by my honorable teacher Asif khan Shakir. I would like to be thankful to my supervisor for his kind support, guidance, and encouragement and I want to express my gratitude towards my parents, teachers, batch mates, and my seniors of DIU for their kind assistance and advice to complete my study.

TABLE OF CONTENTS

Contents

ABSTRACT

Weather Prediction is getting very popular in the current era of Artificial Intelligence. And it is very important to predict rainfall amount in a country like Bangladesh where 50% of people is a farmer. Because the farming in Bangladesh is based on Natural Monsoon. Farmers need to have an idea of how much rain may fall in the coming days. Also, rainfall information helps people become aware of several potential natural disasters. Rainfall prediction is getting very important for Bangladesh's economy day by day. Light or heavy - both rainfalls affect rural and urban life. This work aims to find out the pattern of the average rainfall of Dhaka per month. I used different machine-learning algorithms to predict the future rainfall amount of Dhaka such as Simple linear regression, Multivariate linear regression, Polynomial regression. Because rainfall amount depends on several weather attributes, in linear regression and polynomial regression, we used different attributes to find out which attribute gives the best result. And for multivariate linear regression, we used all possible attribute that has a relation with rainfall.

Keywords: Simple Linear Regression, Polynomial Regression, Multivariate Linear Regression.

CHAPTER 1
INTRODUCTION

1.1. INTRODUCTION

Rainfall is one of the finest gifts of nature and many countries depend on it for agriculture. Rainfall Prediction is a kind of Weather Prediction. In recent years, the rainfall patterns around Bangladesh change fast and continuously. Accurate rainfall prediction is important in daily life. It has a great impact on agriculture as well as natural disaster management.

Cultivation in Bangladesh mainly depends on rainfall and it is important to predict whether it will rain or not. Rainfall prediction can help farmers to make their decision on what to cultivate. Rainfall forecasts help not only farmers but all types of people by informing them of impending natural disasters. The rainfall prediction is important because it's related to landslides, floods, earthquakes. This sort of disaster affects people for many years.

There are mainly two types of approaches used to predict rainfall. The Empirical approach and the dynamical approach. Analysis of Historical data of rainfall and its relationship to a variety of atmospheric and oceanic variables are used in the Empirical Approach. Mostly used empirical Approach that used for predictions are regression, artificial neural networks, and fuzzy logic, and group method of data handling.

The regression model used previous data for estimating future events. Regression is mostly used in business, the biological sciences, climate prediction. In this paper, we used three parametric regression analyses.

We aimed to build a model that can predict the average rainfall of Dhaka in this paper. The dataset we used in this paper contains the daily maximum temperature, minimum temperature, relative humidity, wind speed, cloud coverage, and bright sunshine from the year 1953 to 2013. We calculated the sum of the values of all attributes for a month and then divided them by 28, 29, 30, 31 according to the month to get the weather information of that particular month. After that, we used machine learning algorithms to predict.

1.2. Research Question

Our Research Question is:

- RQ 1: Which Regression Model Gives the best Result(from Linear Regression, Polynomial Regression, Multivariate Linear Regression) for Predicting Rainfall amount of Dhaka.
- RQ 2: Which Attribute gives better results in Linear Regression and Polynomial Regression.

1.3. Research Objectives

- To create a rainfall amount Prediction Model that can predict the rainfall amount of Dhaka.

1.4. Organization of the Thesis

- Chapter 1: Chapter one produces the introduction of the thesis. Here also describe the research objectives and the research question
- Chapter 2: This chapter describes the background, literature review and demonstrates previous work related to this study.
- Chapter 3: This chapter depicts the whole proposed model and architecture.
- Chapter 4: This chapter presents the experiment and result and evaluation of the study
- Chapter 5: This chapter concludes with future scope and limitations.

CHAPTER 2
LITERATURE REVIEW

2.1. LITERATURE REVIEW

Ashfaq Ali Shafin discussed five machine learning algorithms for predicting yearly and seasonal temperature. He used Linear regression, polynomial regression, support vector regressor, and isotonic regression for building the models. He showed that Isotonic Regression performs better than any other regression on the training dataset. But it performed very poorly in the test dataset[1].

Mizanur et al. utilized a model, delivered for anticipating mean temperature that changed with ground-based watched data in Bangladesh during the time of 1979-2006. For the realization of the model execution, they have used the Climate Research Unit (CRU) data[2].

Holmstrom et al. focused on a method that determines the highest and lowest temperature of the next seven days, given the data of the past couple of days[3]. They used a linear regression model and a variation of a functional linear regression model.

Radhika et al used a classification model to predict climate. They used a Support Vector Machine algorithm to forecast Climate[4].

Gnanasankaran et al. discussed three machine learning algorithms. they showed that multiple linear regression gives a better result than artificial neural networks and support vector machines[5].

Ozlem proposed a model to estimate rainfall amount in Esparto using a data mining process. He used monthly rainfall data. And the relative error of this model was 0.7%[6]

.

Wint et al. discussed Multiple polynomial regression model to predict monthly rainfall. He used 36 years of data and showed that there is a close agreement between the predicted and actual rainfall amount[7].

Navid et al. focused on Multiple Linear Regression. They build an MLR equation that can predict the future rainfall amount of Rajshahi[8].

Mizanur Rahman et al. also worked on the prediction of subcontinental Rainfall. They used both simple linear regression and multiple linear regression to predict summer monsoon rainfall[9].

Nawaraj et al. proposed a classification model. They used the Decision tree, Random forest, and Support Vector Machine. They get 80.58% accuracy on Random forest, 79.42% accuracy on Decision Tree, and 78.05 accuracies on Support vector Machine[10].

Saroj et al. used four neural networks for forecasting. For each neural network, they designed four models using a monthly dataset, quarterly dataset, biannual dataset, and annual dataset. The results show that each model increases accuracy by performing sensitivity analysis and removing useless inputs[11].

Prabakaran et al. discussed a Modified Linear Regression. They used multiple linear regression. After completing one iteration, they checked the error rate. Then they perform 4 more iterations to reduce the error[12].

Retius et al. discussed a multiple regression model. They used linear regression with the combination of time series analysis [13]. Goutami used the data from 1871-1999 to estimate average summer rainfall. She used multiple linear regression to predict the monthly rainfall of India[14].

Nikhil et al. used both Linear and Multiple Linear Regression. They used precipitation, vapor pressure, average temperature, and cloud coverage to estimate the rainfall[15].

Samuel et al. also used multiple linear regression to predict monsoon rainfall. They used outgoing longwave radiations, global temperatures, and sunspots out of Tamil Nadu. They used 110 years of data to predict[16].

2.2. SUMMARY OF MOSTLY RELATED WORK

A few research papers are closely related to this work stated in table 2.2.1.

2.2.1 Table of mostly related work

Authors	Title	Objective	Year
Ashfaq Ali Shafin	Machine Learning Approach to Forecast Average Weather Temperature of Bangladesh	Identify the best regression to predict temperature	2019
N. Gnanasankaran, E. Ramaraj	A Multiple Linear Regression Model To Predict Rainfall Using Indian Meteorological Data	To present a new MLR model for rigid rainfall forecasting	2020
Nikhil Sethi, Dr.Kanwal Garg	Exploiting Data Mining Technique for Rainfall Prediction	To predict rain in the future year by knowing climate factors.	2014

This paper concentrates on finding the best regression model for predicting rainfall and the proper attribute to predict rainfall using simple linear regression and polynomial regression in Machine Learning algorithms.

CHAPTER 3
RESEARCH METHODOLOGY

3.1. METHODOLOGY MODEL

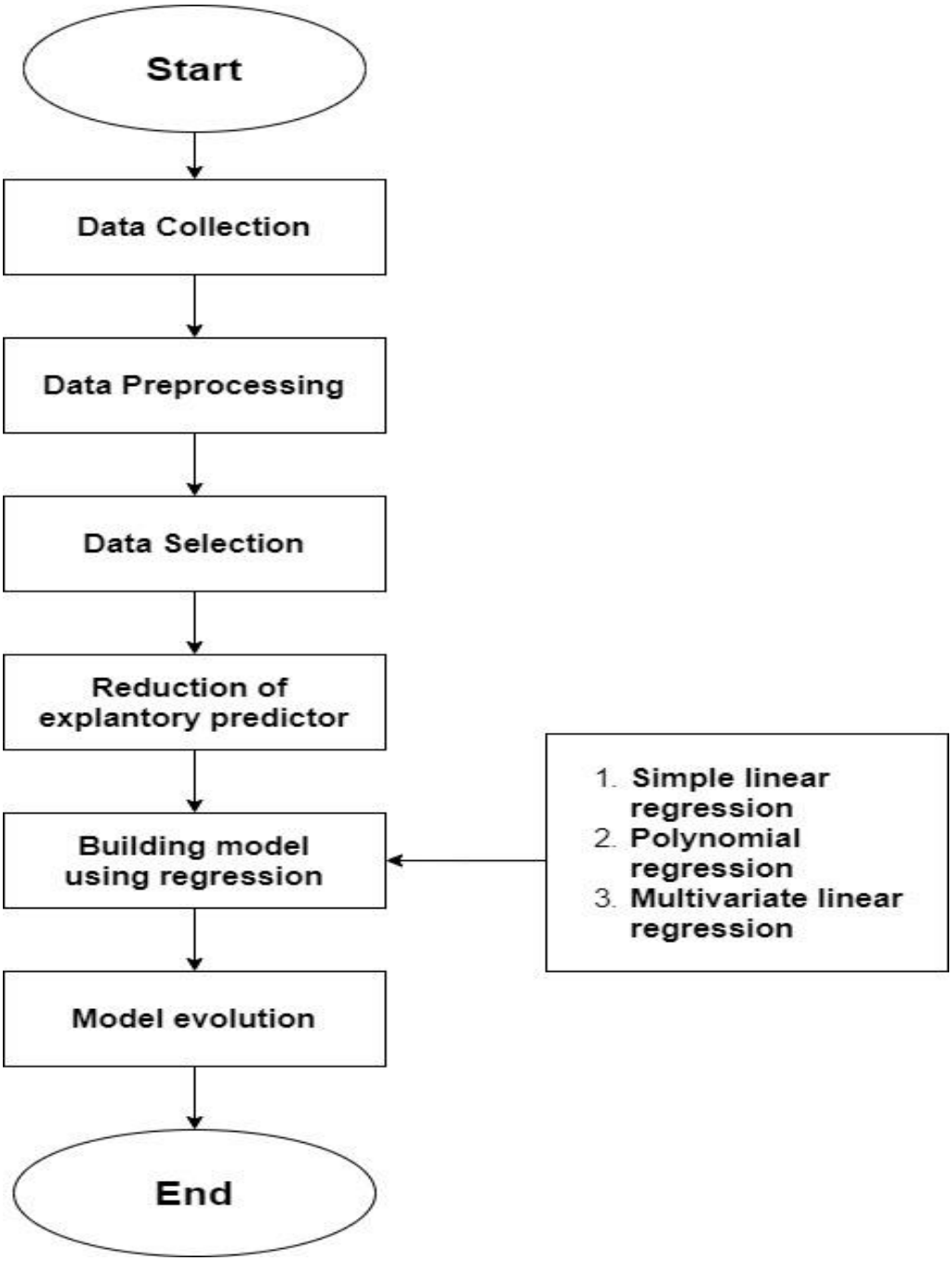


Figure: 3.1.1 Methodology model

3.2. DATASET

We collected our dataset from Kaggle.com. This dataset contains daily maximum temperature, daily minimum temperature, daily relative humidity, daily cloud coverage, daily wind speed, and daily bright sunshine. These six dataset we used as a predictor. We collect daily rainfall data from data.gov.bd. We used this dataset as target data. All the categories data was from 1953-2013. The dataset contains data from numerous weather stations of Bangladesh including Barishal, Bhola, Bogra, Chandpur, Chittagong, Comilla, Dhaka, Dinajpur, and Faridpur. But we have only taken the data related to Dhaka

3.3. PRE-PROCESSING

All our data was daily based. We convert the data into monthly based. For monthly based we have added all days data first, then divided it with 28, 29, 30, 31 based on the respective month. A mathematical equation is given below

$$Monthly\ Data\ x = \sum_{i=1}^{n} Daily\ Data/n \quad \text{.....(1)}$$

Here, n = 28,29,30,31 based on respective month

3.4. DATA VISUALIZATION AND STATISTICS

First, we plot rainfall against corresponding attributes. Here we are showing the relationship between Rainfall with respectively Maximum Temperature, Minimum Temperature, Relative Humidity, wind speed, Cloud Coverage, and Bright sunshine.

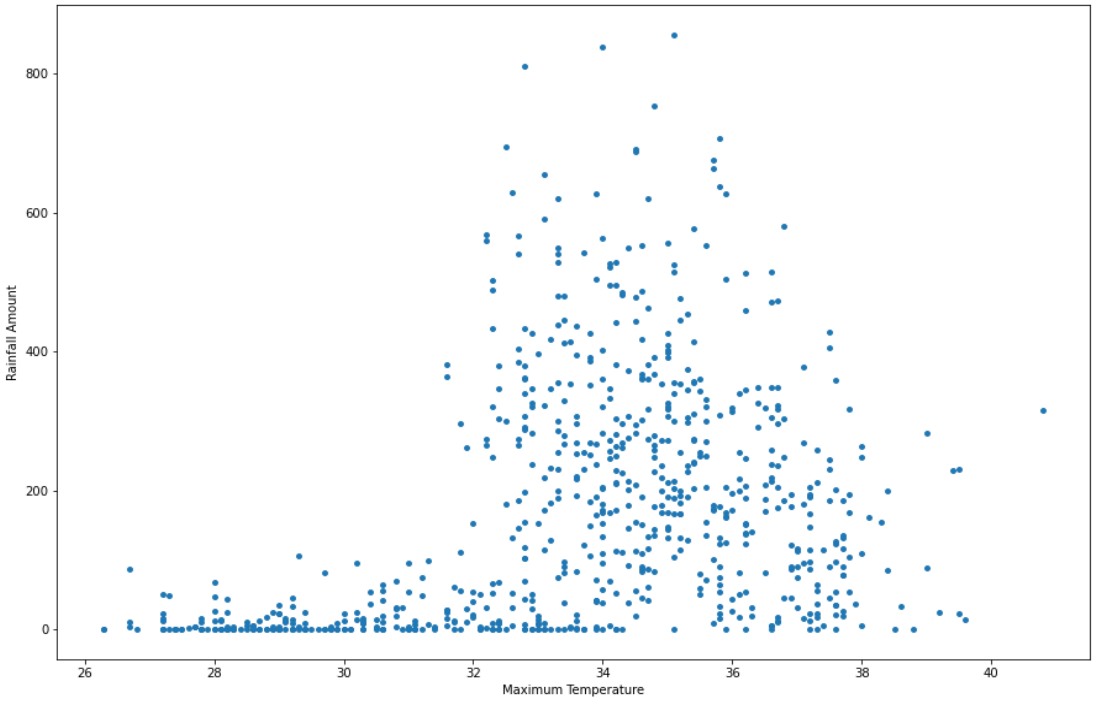


Figure 3.4.1. Relationship between Maximum Temperature and Rainfall

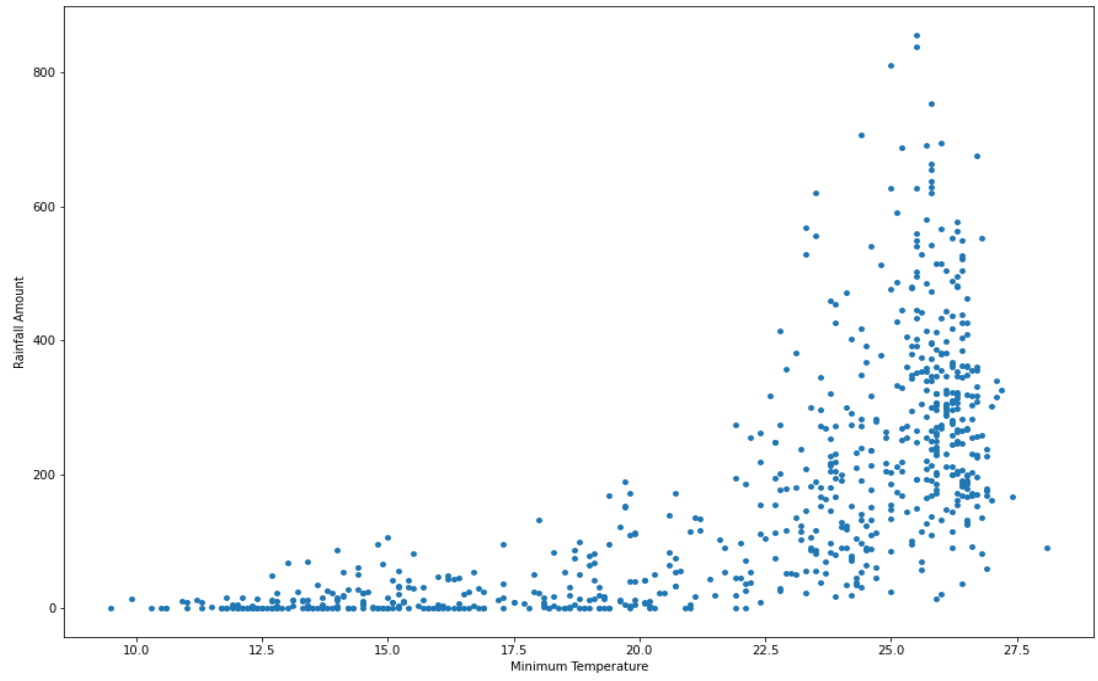


Figure 3.4.2. Relationship between Minimum Temperature and Rainfall

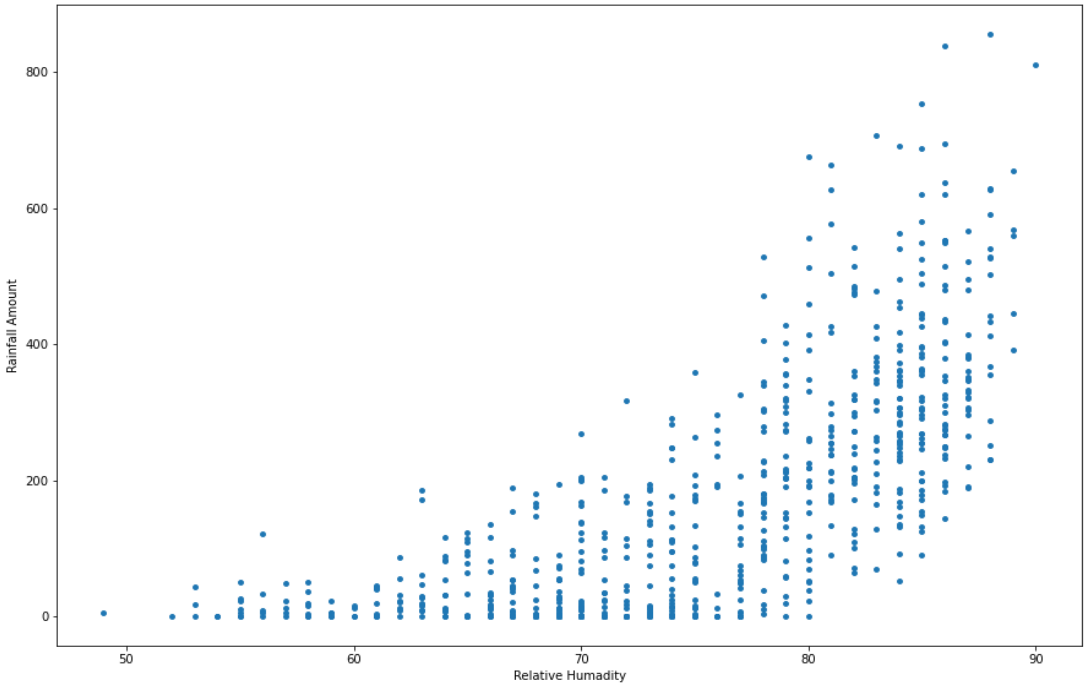


Figure 3.4.3. Relationship between Relative Humidity and Rainfall

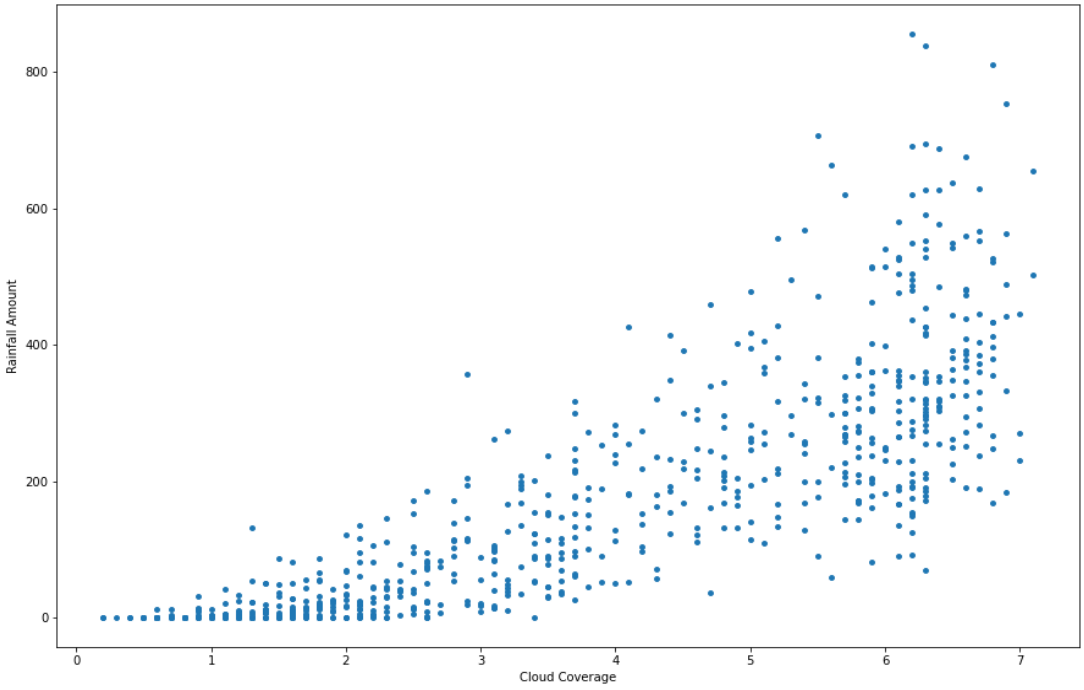


Figure 3.4.4. Relationship between Cloud Coverage and Rainfall

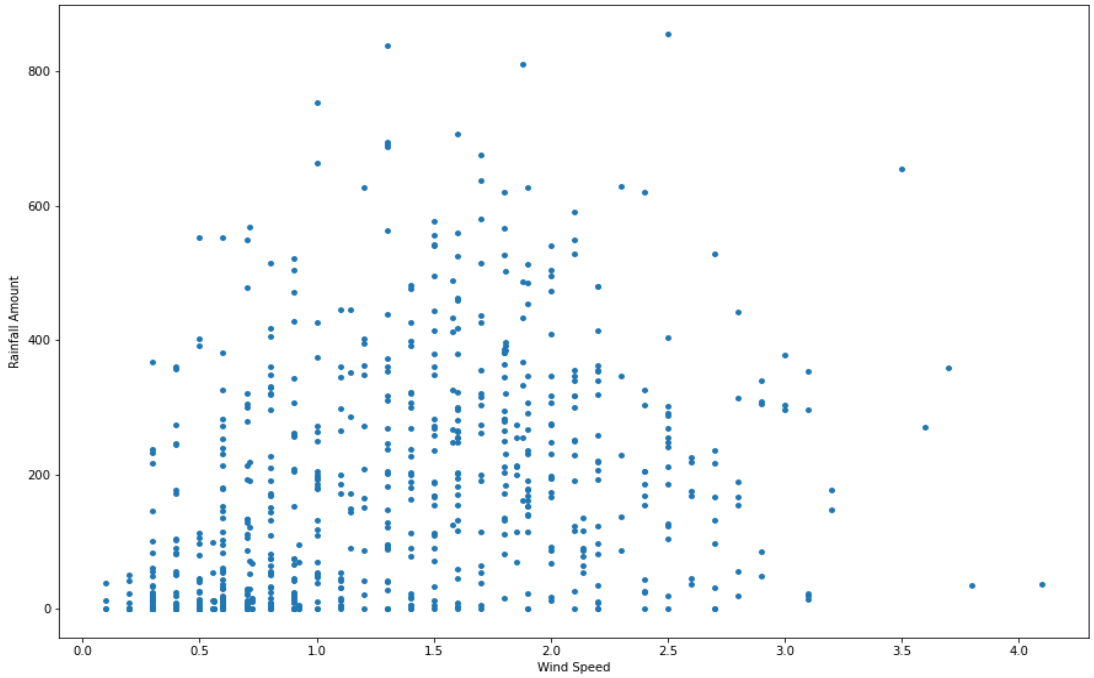


Figure 3.4.5. Relationship between Wind Speed and Rainfall

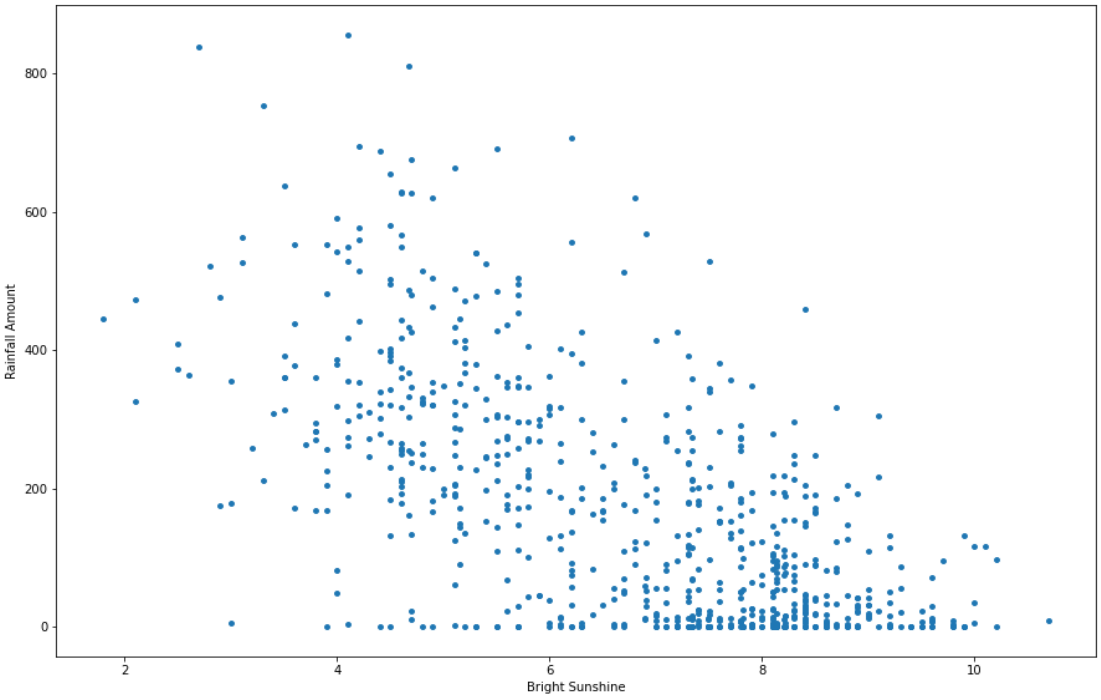


Figure 3.4.6. Relationship between Bright Sunshine and Rainfall

3.5. DATASET SATISTICAL OVERVIEW

3.5.1. Dataset Statistical Overview

	Max Temp	Min Temp	Relative Humidity	Wind Speed	Cloud Coverage	Bright Sunshine	Rainfall
count	732.000000	732.000000	732.000000	732.000000	732.000000	732.000000	732.000000
mean	33.547131	21.470219	75.379781	1.274074	3.766530	6.734277	169.942623
std	2.907880	4.993247	8.592175	0.753721	2.055312	1.741856	174.257148
min	26.300000	9.500000	49.000000	0.100000	0.200000	1.800000	0.000000
25%	32.000000	16.900000	69.000000	0.600000	1.800000	5.300000	13.000000
50%	34.000000	23.650000	76.000000	1.142593	3.600000	7.150000	123.000000
75%	35.600000	25.800000	83.000000	1.807407	5.900000	8.130189	274.500000
max	40.800000	28.100000	90.000000	4.100000	7.100000	10.700000	856.000000

3.6. ESTIMATOR SELECTION

In this study, we have used three machine-learning algorithms. That is described in the table below.

3.6.1. Table of Estimator and their Parameter

Estimator	Parameter
Simple Linear Regression	Default
Polynomial Regression	Degree = 2
Multivariate Linear Regression	Default

This Estimator we used to train our data and predict future Rainfall amounts.

Linear regression is currently the most widely used prediction model. Linear regression is a means of determining the relationship between dependent variables and non-dependent variables. For simple linear regression, assume x is the predictor and independent variable and y is the dependent variable. So the linear regression for x and y is

$$y = w_0 + w_1x + e \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

here, w_0 and w_1 both are weight vectors and e is the error term.

Polynomial regression is a kind of regression analysis in which the relationship between the independent variable x and dependent variable y in n-th degree polynomial in x. Polynomial relapse fits a nonlinear relationship between the worth of x and the corresponding conditional mean of y[1]. In this study, we have used polynomial regression of 2nd-degree. The 2^{nd} –degree polynomial regression equations represented as follows:

$$y = w_0 + w_1x + w_2x^2 + e\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Here, w0, w1, w2 is the weight vector y is output variable rainfall and x is the predictor.

Multivariate Linear regression works just like simple linear regression except it has more than one independent variable. The multivariate linear regression equation is:

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + \ldots\ldots\ldots + w_nx_n \ldots\ldots\ldots(4)$$

To build the multiple linear regression equation the parameter is taken from the training data and variables are extracted from the dataset using correlation.

The linear correlation coefficient measures the strength of the relationship between two variables. It also measures the direction of the relationship between the two variables. Its denoted with r. The mathematical equation for r is given below:

$$r = \frac{n\sum(xiyi) - \sum xi \sum yi}{\sqrt{n\sum(\text{xi}\text{\^{}}2 - (\text{xi})\text{\^{}}2)}\sqrt{n\sum yi^2 - (yi\text{\^{}}2)}} \quad \ldots\ldots(5)$$

Here, n is the sample size, x_i, and y_i are the samples indexed with i.

The coefficient of determination measures how well the regression line represents data if the regression line passes through every point on the scattered plot, it would be ready to explain all of the variations.

CHAPTER 4

RESULT AND DISCUSSION

4.1. RESULT ANALYSIS

In this study, our goal is to make a Rainfall Prediction Model using different types of parameters or attributes.

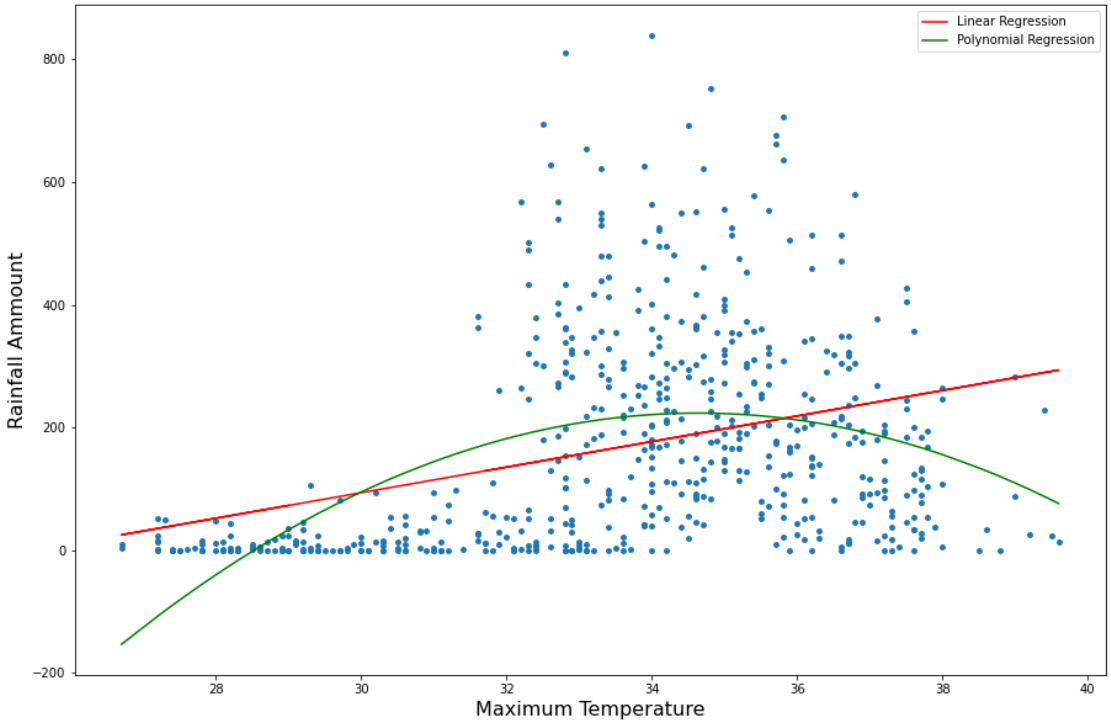


Figure: 4.1.1. Regression Model using Max. Temp. as Predictor

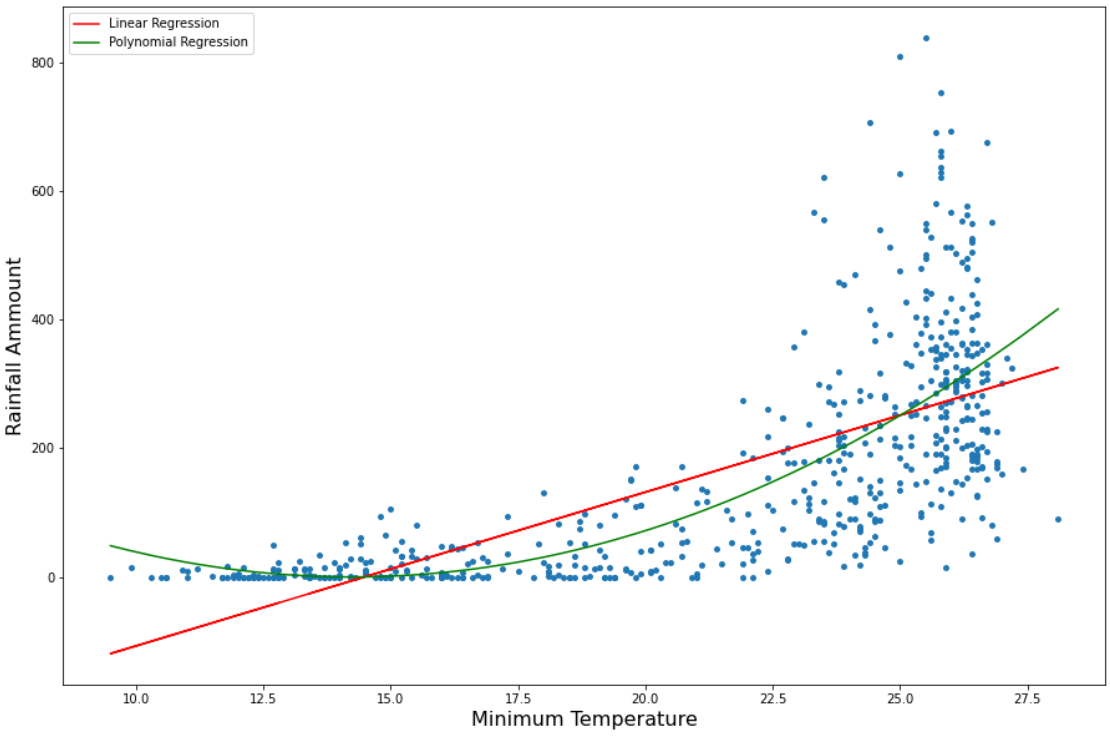


Figure: 4.1.2. Regression Model using Min. Temp. as Predictor

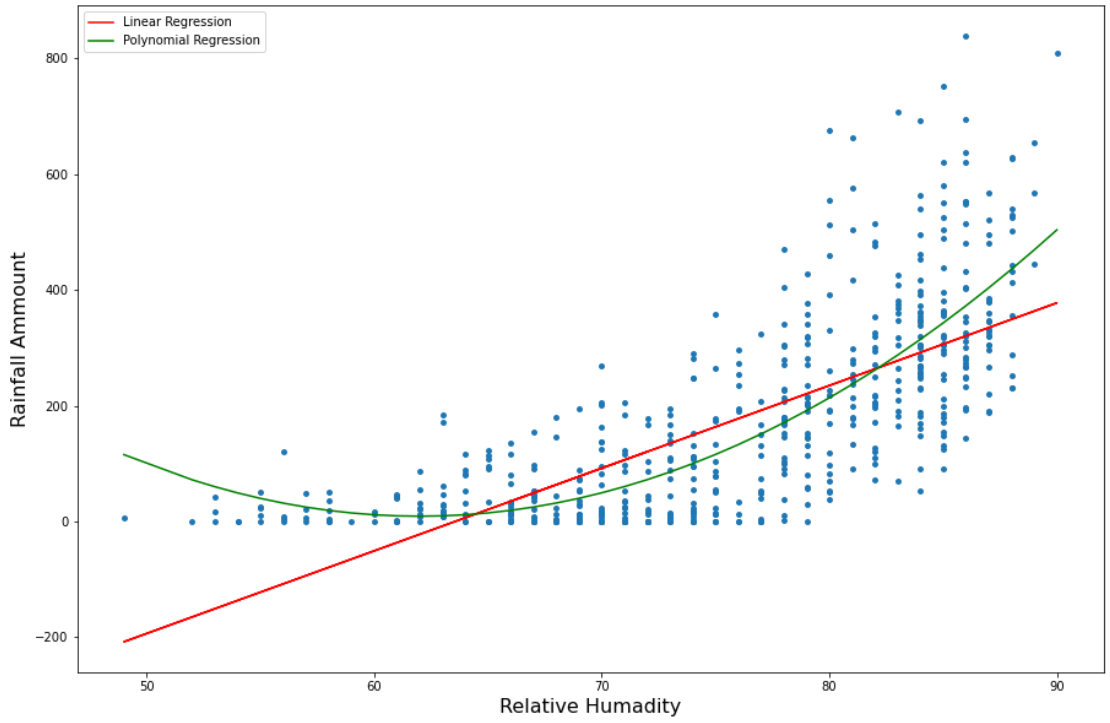


Figure: 4.1.3. Regression Model using Relative Humidity as Predictor

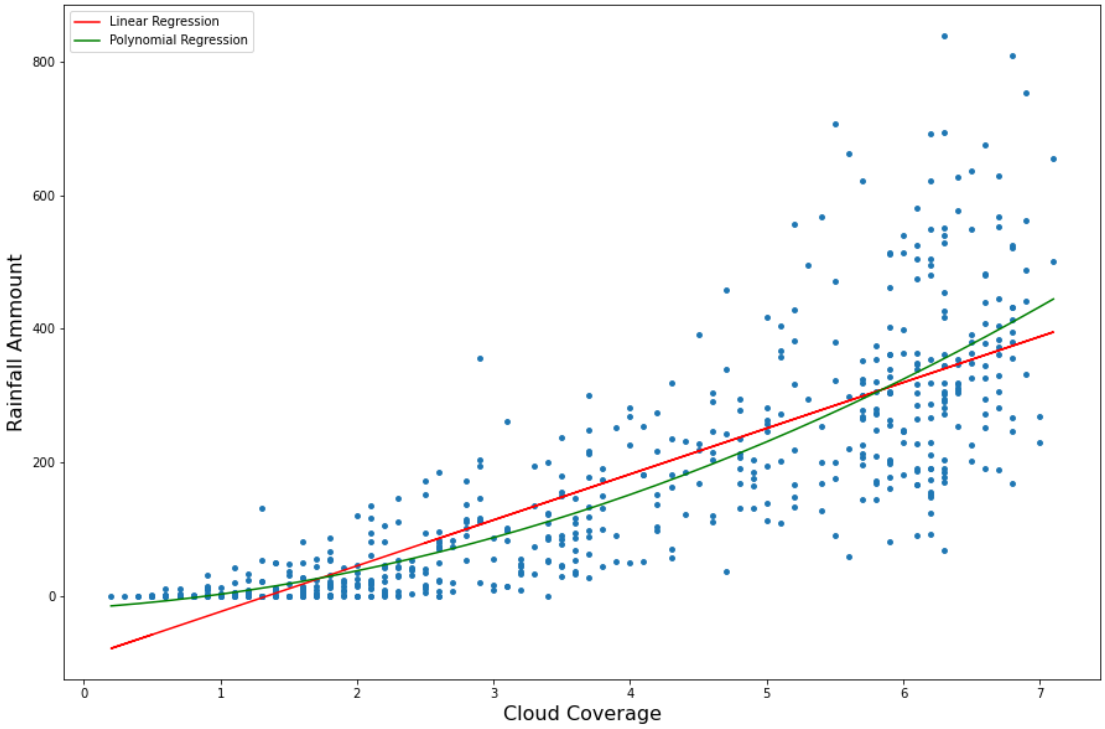


Figure: 4.1.4. Regression Model using Cloud Coverage as Predictor

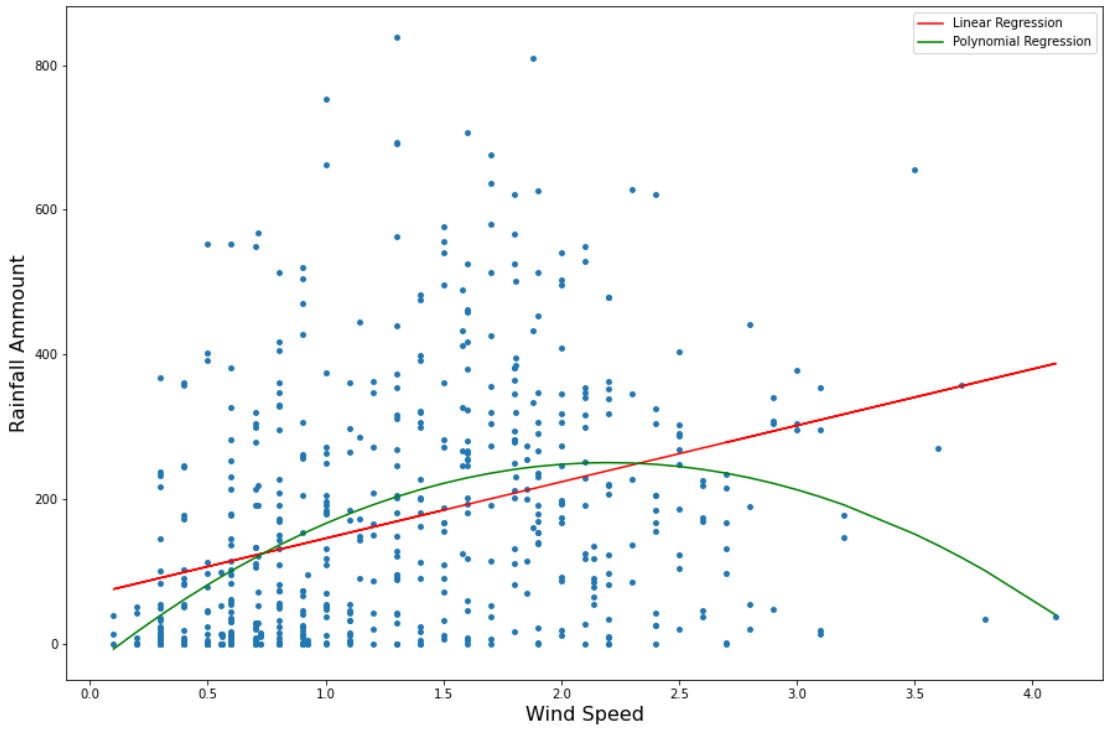


Figure: 4.1.5 Regression Model using Wind Speed as Predictor

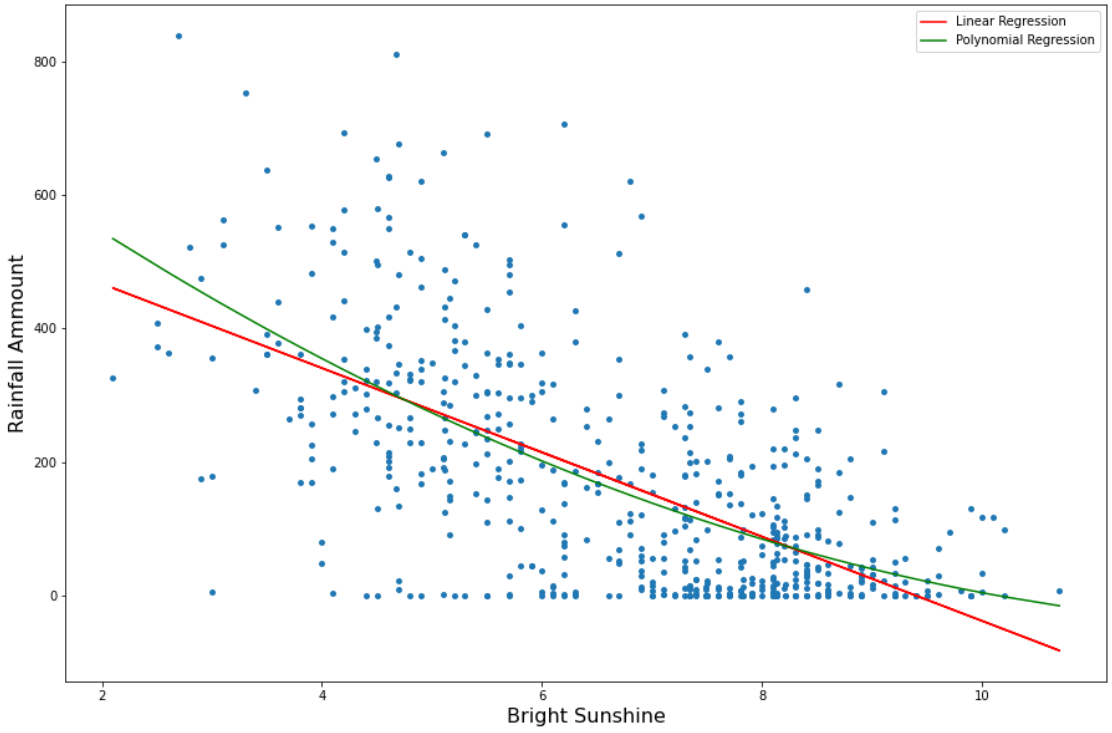


Figure: 4.1.6. Regression Model using Bright Sunshine as Predictor

Figure 4.1.1. represents the monthly rainfall for simple linear and polynomial regression using maximum temperature as the predictor. Both regression models are fitted in the same figure. From the figure, it's clear that Polynomial regression performs better if we use Maximum Temperature as an Independent Variable. Here, for simple linear regression, the coefficient is 20.8139 and the intercept is -503.4897. and for polynomial regression, the coefficient is 414.2479, -5.9799 and the intercept is -6950.5041.

Table 4.1.1: Difference between Test value and Predicted value for Polynomial Regression

Actual Value	Predicted Value	Difference
65	215.476346	-150.476346
590	209.450036	380.549964
0	89.407004	-89.407004
254	218.323905	35.676095
300	222.285846	77.714154

In figure 4.1.2, we used Minimum Temperature as an independent variable and it's clear that the polynomial model gives a better result than simple linear regression. Here, for simple linear regression, the coefficient is 23.9254 and the intercept is -346.559. and for polynomial regression, the coefficient is -61.7318, 2.1684 and the intercept is 439.2778.

Table 4.1.2: Difference between Test value and Predicted value for Polynomial Regression

Actual Value	Predicted Value	Difference
65	87.785546	-22.785546
590	255.924157	334.075843
0	5.235119	-5.235119
254	284.977946	-30.977946
300	210.970618	89.029382

In figure 4.1.3, we used Relative Humidity as an independent variable and we can observe this the polynomial model performs better than simple linear regression. Here, for simple linear regression, the coefficient is 14.2797 and the intercept is -907.4686. and for polynomial regression, the coefficient is -78.1627, 0.6304 and the intercept is 2431.7539.

Table 4.1.3: Difference between Test value and Predicted value for Polynomial Regression

Actual Value	Predicted Value	Difference
65	14.853623	50.146377
590	435.685379	154.314621
0	100.089481	-100.089481
254	342.966929	-88.966929
300	191.574646	108.425354

In figure 4.1.4 we used Cloud Coverage as an independent variable as a predictor and from the figure, it's not cleared which model performs better. Here, for simple linear regression, the coefficient is 68.6058 and the intercept is -91.7909. and for polynomial regression, the coefficient is 13.0438, 7.3304 and the intercept is -17.3991.

Table 4.1.4: Difference between Test value and Predicted value for Polynomial Regression

Actual Value	Predicted Value	Difference
65	76.593788	-11.593788
590	355.720023	234.279977
0	8.809214	-8.809214
254	355.720023	-101.720023
300	189.738392	110.261608

In figure 4.1.5 we used the Wind Speed as an Independent Variable from observation we can see, the polynomial model performs better than simple linear regression. Here, for simple linear regression, the coefficient is 78.0198 and the intercept is 67.6343. and for polynomial regression, the coefficient is 257.068, -58.411 and the intercept is -32.5188.

Table 4.1.5: Difference between Test value and Predicted value for Polynomial Regression

Actual Value	Predicted Value	Difference
65	235.688533	-170.688533
590	249.730775	340.269225
0	92.268589	-92.268589
254	244.306748	9.693252
300	229.257414	70.742586

In figure 4.1.6 we used the Bright Sunshine as an independent variable and it is not clear from the figure that which model performs better. Here, for simple linear regression, the coefficient is -63.1355 and the intercept is 593.2912. and for polynomial regression, the coefficient is -122.4389, 4.5698, and the intercept is 771.6526.

Table 4.1.6: Difference between Test value and Predicted value for Polynomial Regression

Actual Value	Predicted Value	Difference
65	89.587377	-24.587377
590	355.013888	234.986112
0	93.500887	-93.500887
254	299.690969	-45.690969
300	181.663349	118.336651

Now for the Result analysis, we have used four estimators to find out the estimator's outcome. The estimators are, Mean Squared Error, Root Mean Squared Error, Mean absolute error, R2 Score.

The Mean Squared Error is used to measure the quality of an estimator. The Mean Squared Error is the average of the square of the difference between the estimated value and predicted value. The equation to calculate Mean Squared error is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(yi - \hat{y}\text{i})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

The Root Mean Squared Error is the standard deviation of the prediction error. The equation to calculate Root Squared Error is:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - \hat{y}\text{i})2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

The Mean Absolute Error measures the common value of the errors in a fixed of forecasts, without thinking about their direction. It measures accuracy for consecutive variables. The equation is given below:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|yi - \hat{y}\text{i}| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

The R-Squared is used to find out how well the regression model fits the observed data. R2 values range is 0.0-1.0. The higher value of R2 means the model is good. The equation is:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}-\bar{y})2}{\sum_{i=1}^{n}(y-\bar{y})2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

4.2. Performance Table

Table 4.2.1 performance table for figure 4.1.1

	MSE	RMSE	MAE	R2 Score
Simple Linear Regression	26728.07	163.49	125.84	0.16
Polynomial Regression	25908.55	160.96	125.98	0.19

Table 4.2.2 performance table for figure 4.1.2

	MSE	RMSE	MAE	R2 Score
Simple Linear Regression	16174.23	127.18	92.88	0.49
Polynomial Regression	15436.31	124.24	81.81	0.52

Table 4.2.3. Performance table for figure 4.1.3

	MSE	RMSE	MAE	R2 Score
Simple Linear Regression	16127.45	126.99	98.92	0.49
Polynomial Regression	15436.31	124.24	81.81	0.52

Table 4.2.4. Performace table for figure 4.1.4

	MSE	RMSE	MAE	R2 Score
Simple Linear Regression	9898.21	99.49	67.90	0.69
Polynomial Regression	9950.38	99.75	64.08	0.69

Table 4.2.5. Performance table for figure 4.1.5

	MSE	RMSE	MAE	R2 Score
Simple Linear Regression	25434.06	159.48	124.78	0.21
Polynomial Regression	21738.86	147.44	110.20	0.32

Table 4.2.6. Performance table for figure 4.1.6

	MSE	RMSE	MAE	R2 Score
Simple Linear Regression	18310.13	135.31	104.42	0.43
Polynomial Regression	18525.27	136.107	105.40	0.42

From the data in all the tables above, we observed that polynomial regression worked well with most of the independent variables. Simple linear regression has yielded good results using only Bright Sunshine as an independent variable. As, Mean Squared Error, Root Mean Squared Error, Mean Absolute Error are the lower value means better performance and Higher R-squared value means better performance for the model. We can see that Cloud Coverage as an Independent Variable, Polynomial Regression performs better than the simple linear model.

4.3. Multivariate Linear Regression

We have used six independent variables for multiple linear regression. The variables are Maximum Temperature, Minimum Temperature, Relative Humidity, Cloud Coverage, Wind Speed, and Bright Sunshine.

We performed this experiment to find out the accuracy of rainfall prediction using Multivariate Linear Regression.

The Heatmap below showed the relation between all 7 variable:

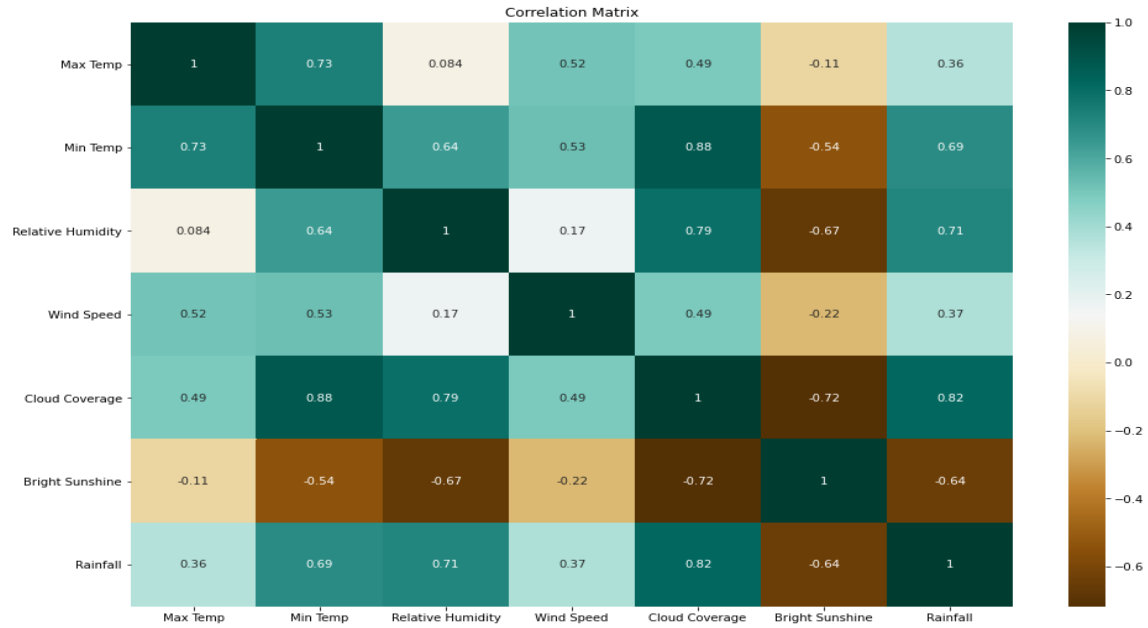


Figure 4.3.1. Correlation Matrix

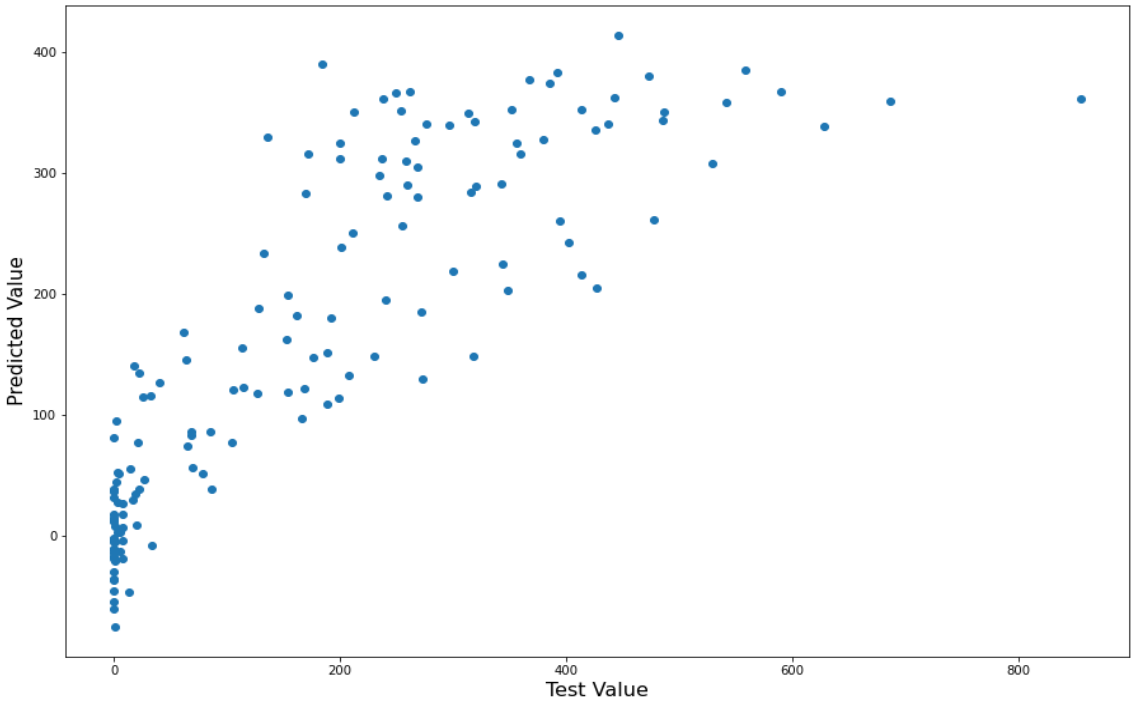


Figure 4.3.2 Relationship Between Test and Predicted value

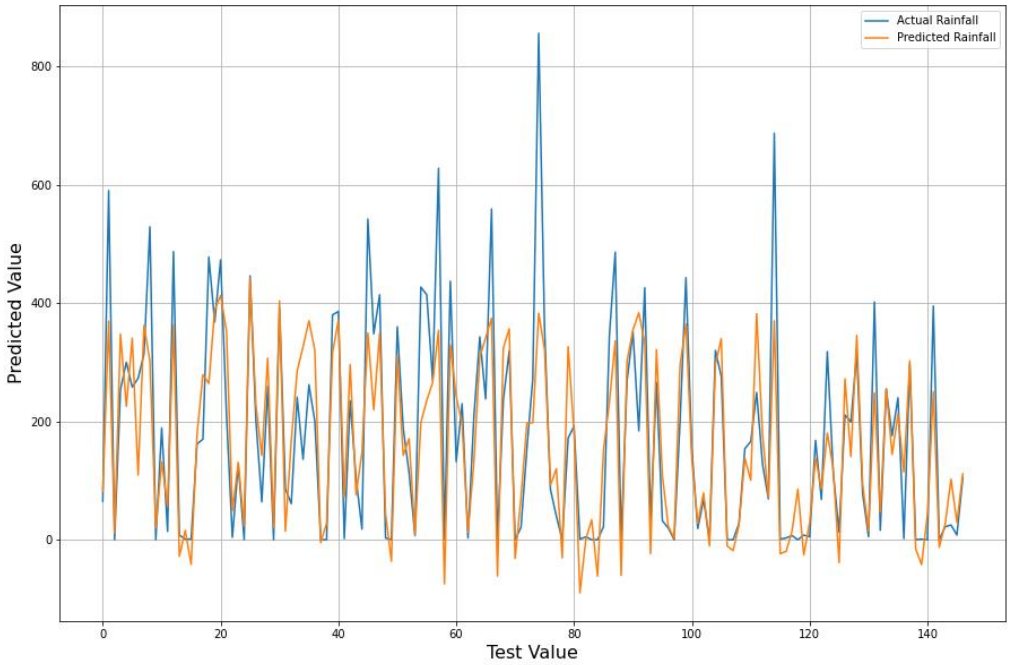


Figure 4.3.3: Relationship Between Test and Predicted value

Figure 4.3.2 and Figure 4.3.3 Represent the relationship between the Test value with the value predicted using Multivariate Linear Regression.

Table 4.3.1. Correlation of Independent Variable with Rainfall

No	Independent Variable	Coefficient
1	Max Temp.	11.99898205
2	MinTemp.	-10.55965366
3	Relative Humidity	5.12475019
4	Cloud Coverage	1.62941764
5	Wind Speed	59.28880261
6	Bright Sunshine	-10.8142654

Table 4.3.2. Performance table for figure 4.1.6

	MSE	RMSE	MAE	R2 Score
Multivariate Linear Regression	9250.8183	96.1812	66.57	0.71

Table 4.3.3: Difference between Test value and Predicted value

Actual Value	Predicted Value	Difference
65	81.249613	-16.249613
590	369.541386	220.458614
0	11.374422	-11.374422
254	347.448544	-93.448544
300	225.568888	74.431112

CHAPTER 5

CONCLUSION

5.1. FINDINGS AND CONTRIBUTIONS

My goal for this study was to create the best machine learning model for predicting rainfall in Dhaka and to find out the best weather attribute with which rainfall is well related and which gives a better prediction of rainfall. And we find out that Multivariate Linear regression Gives the best result among the three regressions that I discussed in this paper. And I also find out that if we use simple linear regression or polynomial regression Cloud coverage will be the best predictor or independent variable to predict the rainfall amount of Dhaka.

5.2. FUTURE WORK

Rainfall amount prediction let alone not be very useful to forecasting the weather of a country. So, in the future, we want to forecast the weather in Bangladesh. Where we will forecast all the weather attribute like Maximum Temperature, Minimum Temperature, Humidity, and outlook prediction.

We expect that study will help to grow our country's economy and lifestyle

REFERENCES

[1] Ashfaq Ali Shafin, 2019. Machine Learning Approach to Forecast Average Weather Temperature of Bangladesh

[2] Md. Mizanur Rahman, Nazlee Ferdousi, 2011. Rainfall and temperature scenario for Bangladesh using 20 km mesh AGCM

[3] Mark Holmstrom, Dylan Liu, Christopher Vo, 2016. Machine Learning Applied to Weather Forecasting

[4] Y.Radhika, M.Shashi, 2009. Atmospheric Temperature Prediction using Support Vector Machines

[5] N. Gnanasankaran, E. Ramaraj, 2020. A Multiple Linear Regression Model To Predict Rainfall Using Indian Meteorological Data

[6] Ozlem Terzi, 2012. Monthly Rainfall Estimation Using Data-Mining Process

[7] Wint Thida Zaw and Thinn Thu Naing, 2008. Empirical Statistical Modeling of Rainfall Prediction over Myanmar

[8] MAI Navid, NH Niloy, 2018. Multiple Linear Regressions for Predicting Rainfall for Bangladesh

[9] Md Mizanur Rahman, M Rafiuddin, Md Mahbub Alam, 2012. Seasonal forecasting of Bangladesh summer monsoon rainfall using simple multiple regression model

[10] Nawaraj Paudel and Tekendra Nath Yogi, 2020. COMPARATIVE STUDY OF MACHINE LEARNING ALGORITHMS FOR RAINFALL PREDICTION – A CASE STUDY IN NEPAL

[11] Saroj Kr. Biswas*, Leniency Marbaniang, Biswajit Purkayastha, Manomita Chakraborty, Heisnam Rohen Singh, and Monali Bordoloi, 2016. Rainfall forecasting by relevant attributes using artificial neural networks – a comparative study

[12] S. Prabakaran, P. Naveen Kumar and P. Sai Mani Tarun, 2017. RAINFALL PREDICTION USING MODIFIED LINEAR REGRESSION

[13] Retius Chifurira, 2014. A Weighted Multiple Regression Model to Predict Rainfall Patterns: Principal Component Analysis approach

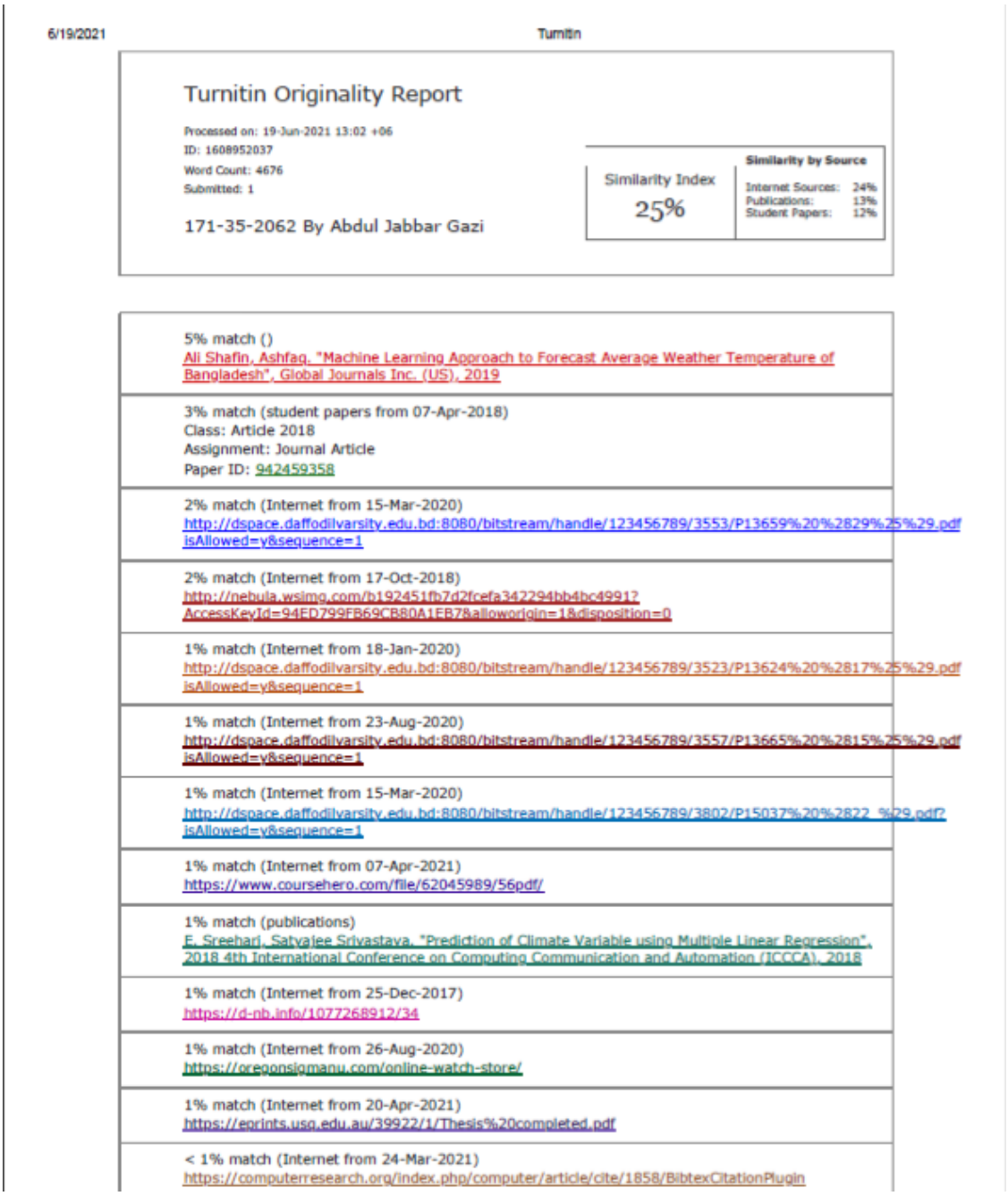
[14] Goutami Bandyopadhyay, 2017. The Prediction of Indian Monsoon Rainfall: A Regression Approach

[15] Nikhil Sethi, Dr.Kanwal Garg, 2014. Exploiting Data Mining Technique for Rainfall Prediction

[16] R. Samuel Selvaraj, Raajalakshmi Aditya, 2011. Statistical Method of Predicting the Northeast Rainfall of Tamil Nadu

[17] Md. Didarul Islam Bhuyan, Md. Mohymenul Islam, Md. Ebrahim Khalil Bhuiyan, 2018. A Trend Analysis of Temperature and Rainfall to Predict Climate Change for Northwestern Region of Bangladesh

PLAGARISM REPORT



6/19/2021 Turnitin

Turnitin Originality Report

Processed on: 19-Jun-2021 13:02 +06
ID: 1608952037
Word Count: 4676
Submitted: 1

171-35-2062 By Abdul Jabbar Gazi

Similarity Index	Similarity by Source
25%	Internet Sources: 24% Publications: 13% Student Papers: 12%

5% match ()
Ali Shafin, Ashfaq. "Machine Learning Approach to Forecast Average Weather Temperature of Bangladesh", Global Journals Inc. (US), 2019

3% match (student papers from 07-Apr-2018)
Class: Article 2018
Assignment: Journal Article
Paper ID: 942459358

2% match (Internet from 15-Mar-2020)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/3553/P13659%20%2829%25%29.pdf?isAllowed=y&sequence=1

2% match (Internet from 17-Oct-2018)
http://nebula.wsimg.com/b192451fb7d2fcefa342294bb4bc4991?AccessKeyId=94ED799FB69CB80A1EB7&allowoorigin=1&disposition=0

1% match (Internet from 18-Jan-2020)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/3523/P13624%20%2817%25%29.pdf?isAllowed=y&sequence=1

1% match (Internet from 23-Aug-2020)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/3557/P13665%20%2815%25%29.pdf?isAllowed=y&sequence=1

1% match (Internet from 15-Mar-2020)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/3802/P15037%20%2822_%29.pdf?isAllowed=y&sequence=1

1% match (Internet from 07-Apr-2021)
https://www.coursehero.com/file/62045989/56pdf/

1% match (publications)
E. Sreehari, Satyajee Srivastava. "Prediction of Climate Variable using Multiple Linear Regression", 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018

1% match (Internet from 25-Dec-2017)
https://d-nb.info/1077268912/34

1% match (Internet from 26-Aug-2020)
https://oregonsigmanu.com/online-watch-store/

1% match (Internet from 20-Apr-2021)
https://eprints.usq.edu.au/39922/1/Thesis%20completed.pdf

< 1% match (Internet from 24-Mar-2021)
https://computerresearch.org/index.php/computer/article/cite/1858/BibtexCitationPlugin