supervised by: Dr. Safaa Alsfry

Section : I1

Layal Soud Halwani
2007896

Afnan Tariq Algogandi
2007926

Kaltham Mohsen Alshayeb
1915435

PROJEC

TEMPERATURE PREDICTION

CCAI-312: PATTERN RECOGNITION

INTRODUCTION

When we talk about the weather, it's an essential thing in our lives, but it can surprise us. Sometimes we may not even know that something is coming that could put us in danger ، like forest fires or snowstorms.
Therefore, it would be useful to be able to predict the weather and its fluctuations before it happens to beware and prepare for it.

So, we developed a model that can predict air temperature according to atmospheric pressure.

PROBLEM DESCRIPTION

In high temperatures, we have to drink enough water and not be in the sun for a long time so that we don't get sunstroke, and in cold temperatures wear warm clothes, etc.
But how can we know what to do if we cannot know the air temperature?

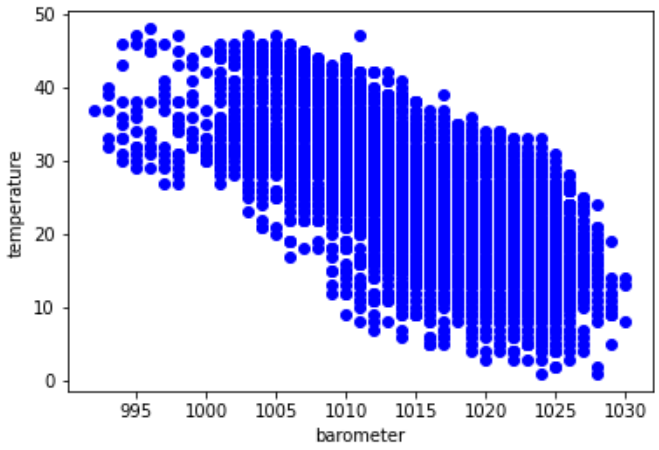
So we tried to solve this problem to help humans to adapt to life on earth by predicting the temperature.

In this project, we predict the temperature by taking the atmosphere as an input and giving the temperature as an output.

DATA DESCRIPTION

In this project we used "Saudi Arabia weather history" dataset from Kaggle.
We took 2 attributes (temp, barometer), type(int64, float64), and 5000 random examples, 80% for training and 20% for testing.

VISUALIZING DATA:



	temp	barometer
count	5000.000000	5000.000000
mean	24.597000	1015.515400
std	8.760225	7.035287
min	1.000000	992.000000
25%	18.000000	1011.000000
50%	24.000000	1016.000000
75%	31.000000	1021.000000
max	48.000000	1030.000000

	temp	barometer
114855	24	1011.0
130335	23	1018.0
83787	24	1021.0
94971	25	1015.0
200126	16	1022.0
...	...	...
89394	32	1010.0
160523	22	1006.0
102897	34	1011.0
15851	15	1020.0
64662	35	998.0

5000 rows × 2 columns

METHOD

We used 3 different models:

1. Linear Regression
2. Batch Gradient Descent
3. Polynomial Regression

Tried different polynomial degrees, learning rates, and number of iterations.
We compared them by using:

1. Mean absolute error (MAE)
2. Mean sum of squares error (MSSE)
3. Root mean square error (RMSE)

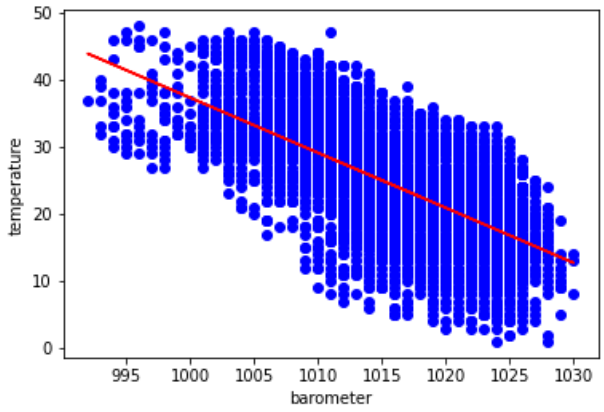
CORRELATION

Y = W1 X + W0
Temp = coefficients * barometer + intercept

Data Description & Method

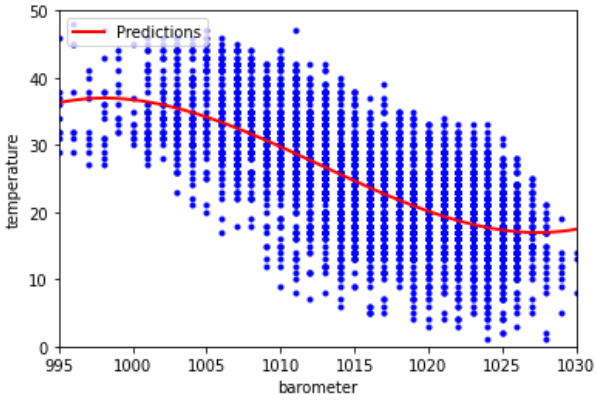
EXPERIMENT AND RESULTS

1 LINEAR REGRESSION



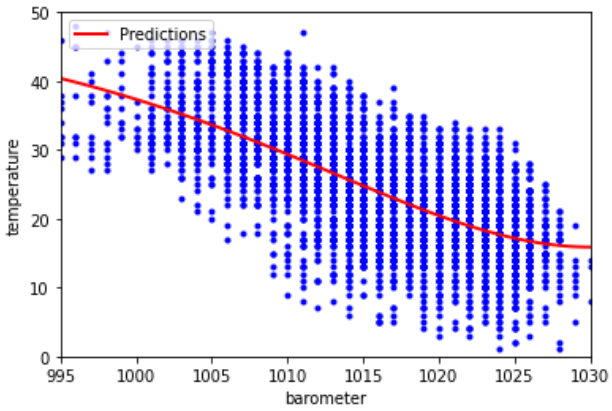
1. MEAN ABSOLUTE ERROR : 5.30
2. MEAN SUM OF SQUARES ERROR : 42.99
3. ROOT MEAN SQUARE ERROR: 6.56

2 POLYNOMIAL REGRESSION



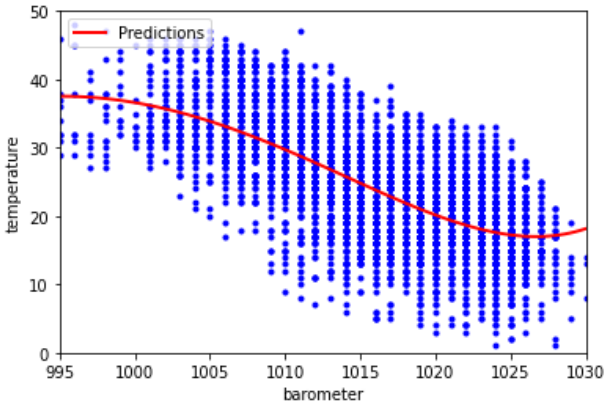
1. MEAN ABSOLUTE ERROR : 5.26
2. MEAN SUM OF SQUARES ERROR : 42.29
3. ROOT MEAN SQUARE ERROR: 6.50

- POLYNOMIAL DEGREE: 3



1. MEAN ABSOLUTE ERROR : 5.28
2. MEAN SUM OF SQUARES ERROR : 42.60
3. ROOT MEAN SQUARE ERROR: 6.53

- POLYNOMIAL DEGREE: 100

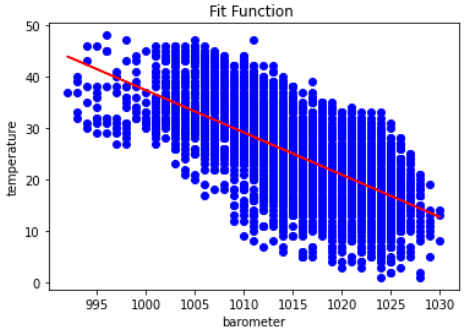


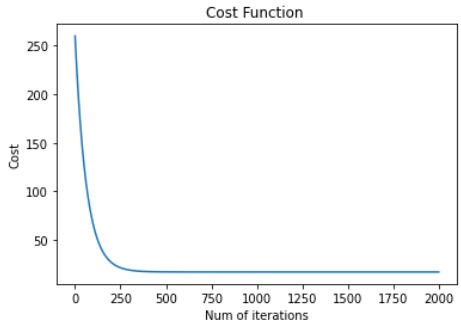
1. MEAN ABSOLUTE ERROR : 5.27
2. MEAN SUM OF SQUARES ERROR : 42.46
3. ROOT MEAN SQUARE ERROR: 6.52

- POLYNOMIAL DEGREE: 30

EXPERIMENT AND RESULTS

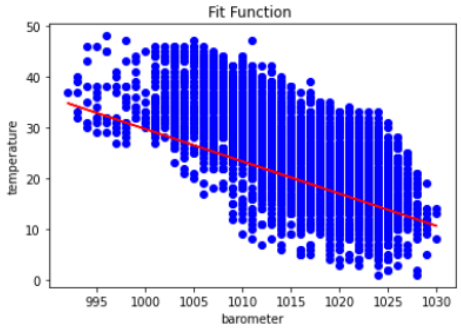
3 BATCH GRADIENT DESCENT

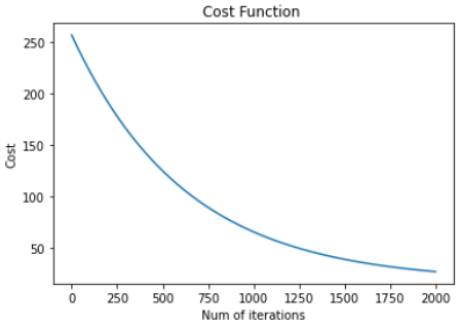




- LEARNING RATE: 0.01
- NUMBER OF ITERATIONS: 2000
- GRADIENT DESCENT: 24.5 , -5.8

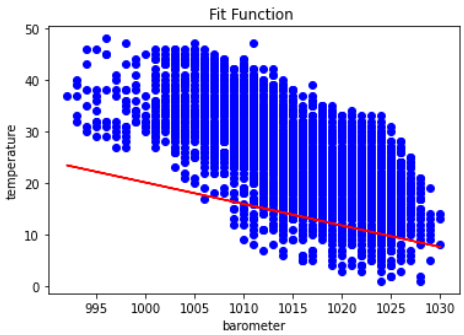
1. MEAN ABSOLUTE ERROR : 5.30
2. MEAN SUM OF SQUARES ERROR : 42.90
3. ROOT MEAN SQUARE ERROR: 6.55

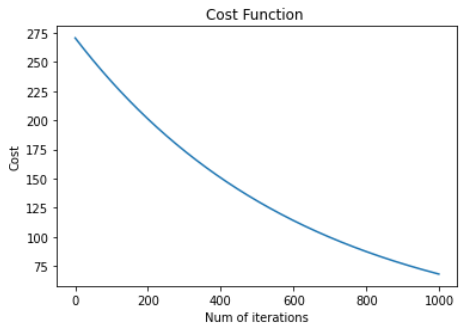




- LEARNING RATE: 0.001
- NUMBER OF ITERATIONS: 2000
- GRADIENT DESCENT: 19.82 ,-4.49

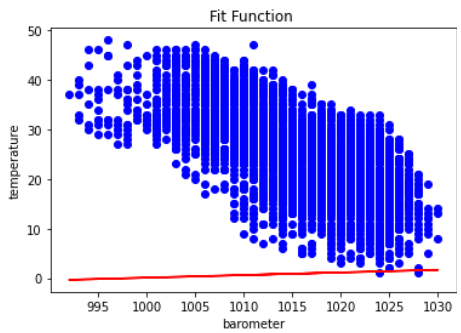
1. MEAN ABSOLUTE ERROR : 6.76
2. MEAN SUM OF SQUARES ERROR : 68.40
3. ROOT MEAN SQUARE ERROR: 8.27

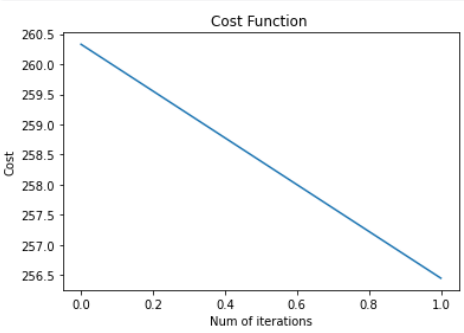




- LEARNING RATE: 0.001
- NUMBER OF ITERATIONS: 1000
- GRADIENT DESCENT: 13.6 , -2.9

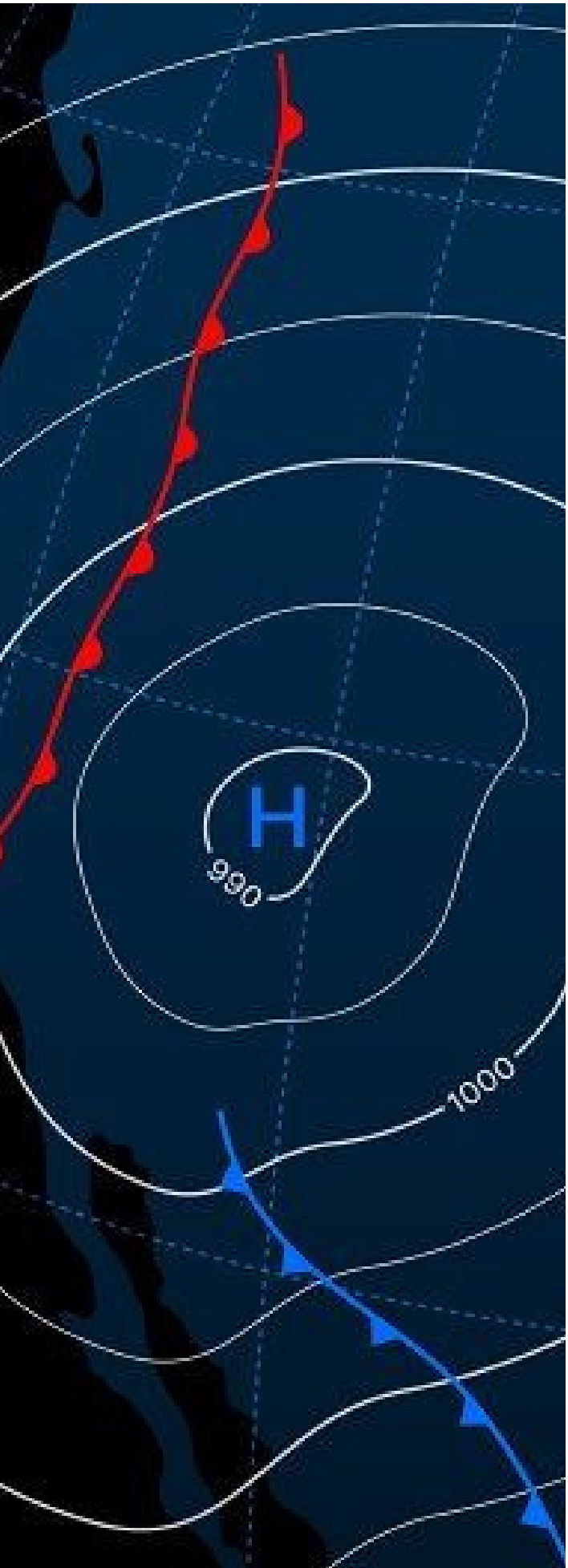
1. MEAN ABSOLUTE ERROR : 11.38
2. MEAN SUM OF SQUARES ERROR : 173.24
3. ROOT MEAN SQUARE ERROR: 13.16





- LEARNING RATE: 0.01
- NUMBER OF ITERATIONS: 2
- GRADIENT DESCENT: 1.0 , 0.3

1. MEAN ABSOLUTE ERROR : 23.54
2. MEAN SUM OF SQUARES ERROR : 637.33
3. ROOT MEAN SQUARE ERROR: 25.25



DISCUSSION

According to the previous results among all types of models, third-order polynomial regression appears to have the best fit with least cost, that is, increasing polynomial order does not imply a greater fitting. And we see in Gradient Descent model how the amount of learning rate affects to the speed of movement toward optimal weights even with the number of iterations fixed, and the plot a cost function showed us how the cost function decreased after each iteration.

CONCLUSION

There is so much information about weather temperatures and their benefits and risks on the environment. Everyone must preserve the environment by predicting the temperature in advance to avoid risks as much as possible. This helps us live in a healthy environment that is suitable for all of us and the next generations. So in this project, we focused on temperature prediction by taking the atmosphere as the input and giving the temperature as the output. And the result was in some models increasing polynomial order imply a greater fitting, While other models do not apply to it.

REFERENCES

Saudi Arabia weather history dataset:
https://www.kaggle.com/datasets/esraamadi/saudi-arabia-weather-history

part of method:
https://www.kaggle.com/code/tentotheminus9/linear-regression-from-scratch-gradient-descent