IoT Based Automated Weather Report Generation and Prediction Using Machine Learning

APOORVA PARASHAR
COMPUTER SCIENCE AND ENGINEERING
MAHARSHI DAYANAND UNIVERSITY *ROHTAK, INDIA*

ANUBHA PARASHAR
COMPUTER SCIENCE AND ENGINEERING
MANIPAL UNIVERSITY JAIPUR

Contents

Introduction

- Weather forecasting is the application of science and technology to predict the state of the atmosphere for a given location.
- Weather is driven by air pressure (temperature and moisture) differences between one place and another. This pressure and temperature differences can occur due to the sun angle at any particular spot, which varies by latitude from the tropics.
- The atmosphere is a chaotic system, so small changes to one part of the system can grow to have large effects on the system as a whole. This makes it difficult to accurately predict weather few days in advance.

About the Project

We aim to create a Weather Monitoring Device which would measure various parameters such as : Temperature , Humidity , Pressure , Rain Fall , PM concentration , CO concentration etc.

These data would then be sent to cloud using wifi module where analysis will be conducted using Matlab.

Our web portal will be showing a graphical presentation of our collected data.

With the help of Machine Learning we aim to forecast the weather of coming days. Air quality for an area will also be displayed.

Literature Survey

1

1. Forecasting power output of photovoltaic systems based on weather classification and support vector machines, J Shi, WJ Lee, Y Liu, Y Yang IEEE Transactions on 2012

2

2 . Predicting weather events using fuzzy rule based system, MSK Awan, MM Awais - Applied Soft Computing, 2015 - Elsevier

3

3. Diagnosis of "forecast analysis" differences of a weather prediction system S Dirren, M Didone, HC Davies - Geophysical research letters, 2016 - Wiley Online Library

4

4. The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation

M Xue, D Wang, J Gao, K Brewster 2016 - Springer

- After reviewing many articles, there are presently no papers that mention monitoring the combination of more features like temperature, lighting and humidity in one integrated system .
- In addition to this, there is one research paper that has discussed monitoring these three environmental conditions; however, there has been no mention about forecasting Temperature using it .
- So our main idea was to coin a system that can sense the main components that formulates the weather and can be able to forecast the weather without human error.

Problem Statement

- Though there are few weather reporting models which already exist but in today's world except for weather, concentration of various components of air and its quality are also gaining importance.

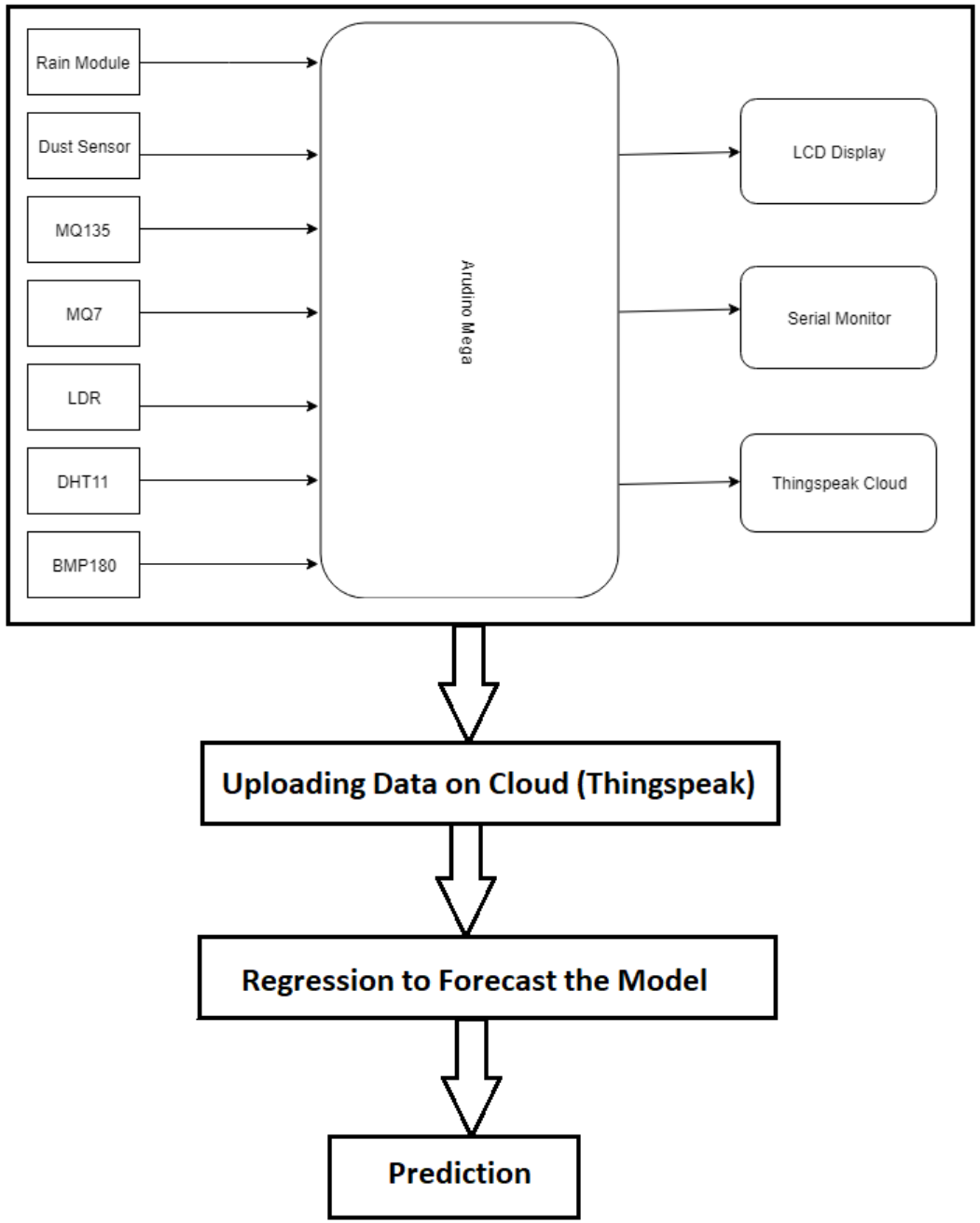
(An operational system for generating impact of air quality in the Australian region and their impact on numerical weather prediction JL Marshall, N Pescod, B Seaman - Nature 2009)

- There is no such model which can forecast weather, determine various components in air and can check or predict air quality.
- Previous weather forecasting models used the complicated blend of mathematical instruments which was insufficient in order to get higher classification rate.
- Using different components to collect and assess data will be expensive.

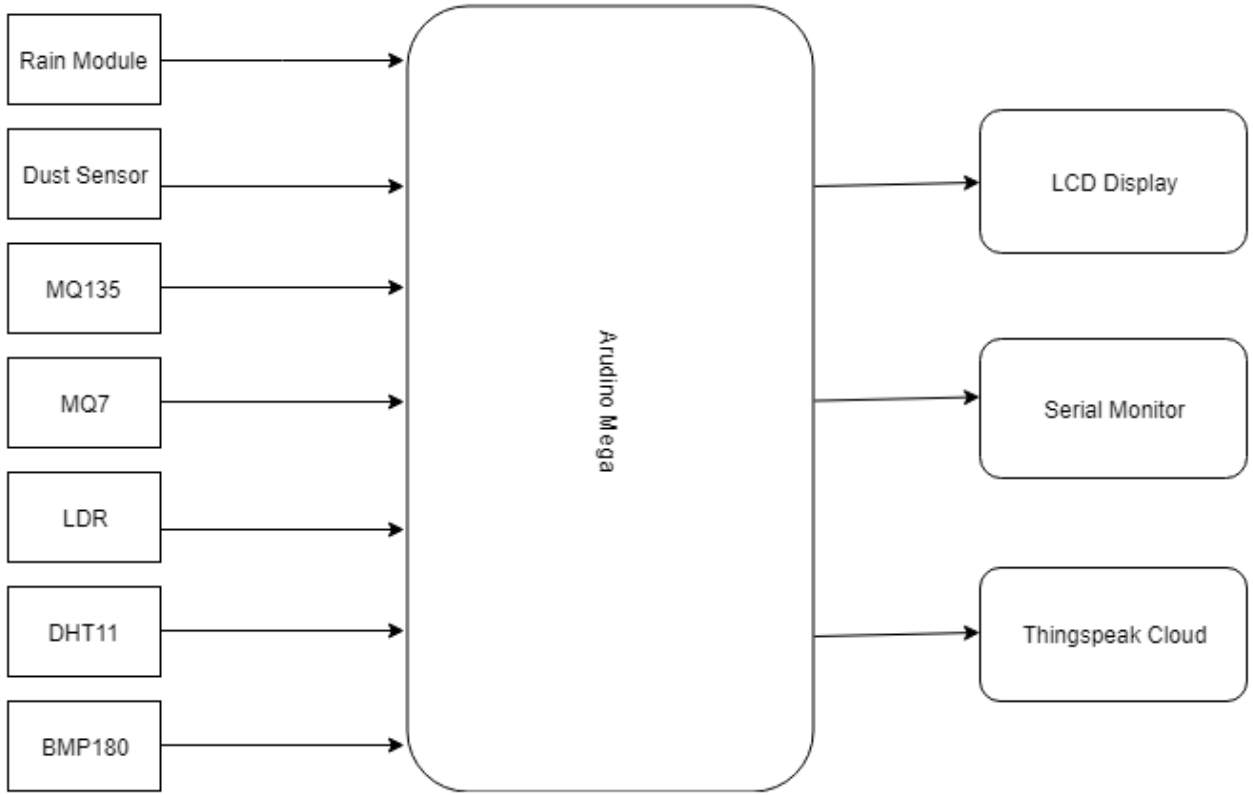
Solution

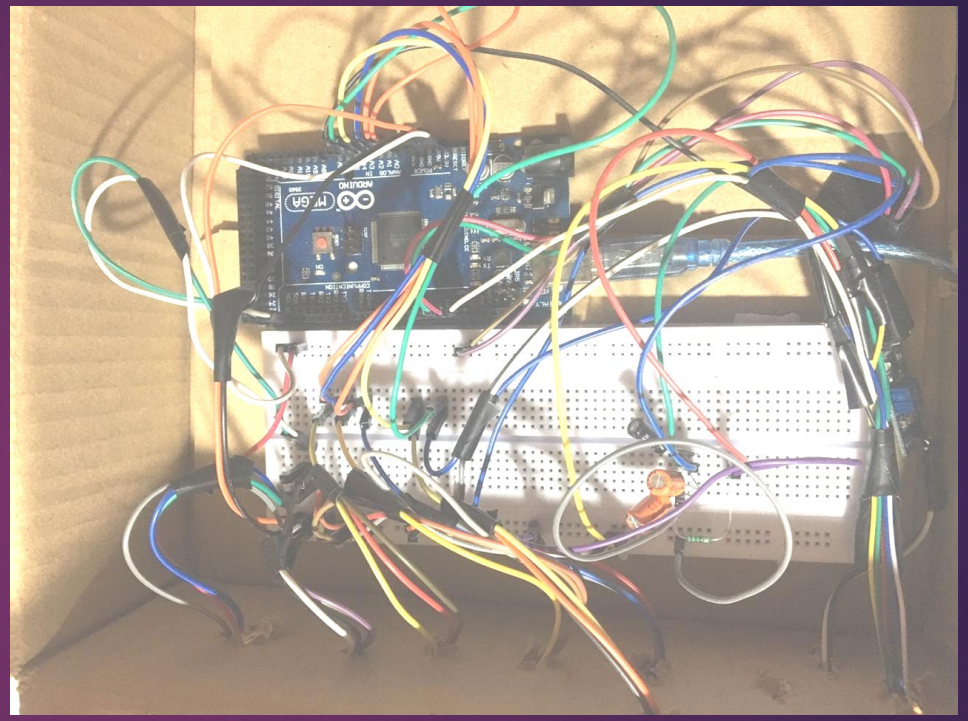
- We in our project aim to integrate various sensors to collect data using which we can forecast weather and calculate air quality as well.
- In our model, we are using more sensors and hence are generating and assessing more parameters.
- Machine learning algorithms will be used which will help us to predict the weather of the coming days based on the historical data of that location hence aiming at higher accuracy at lower cost.
- Also it will be the first setup which will monitor air quality along with forecasting weather and present it in real time.

Methodology



Architecture Design





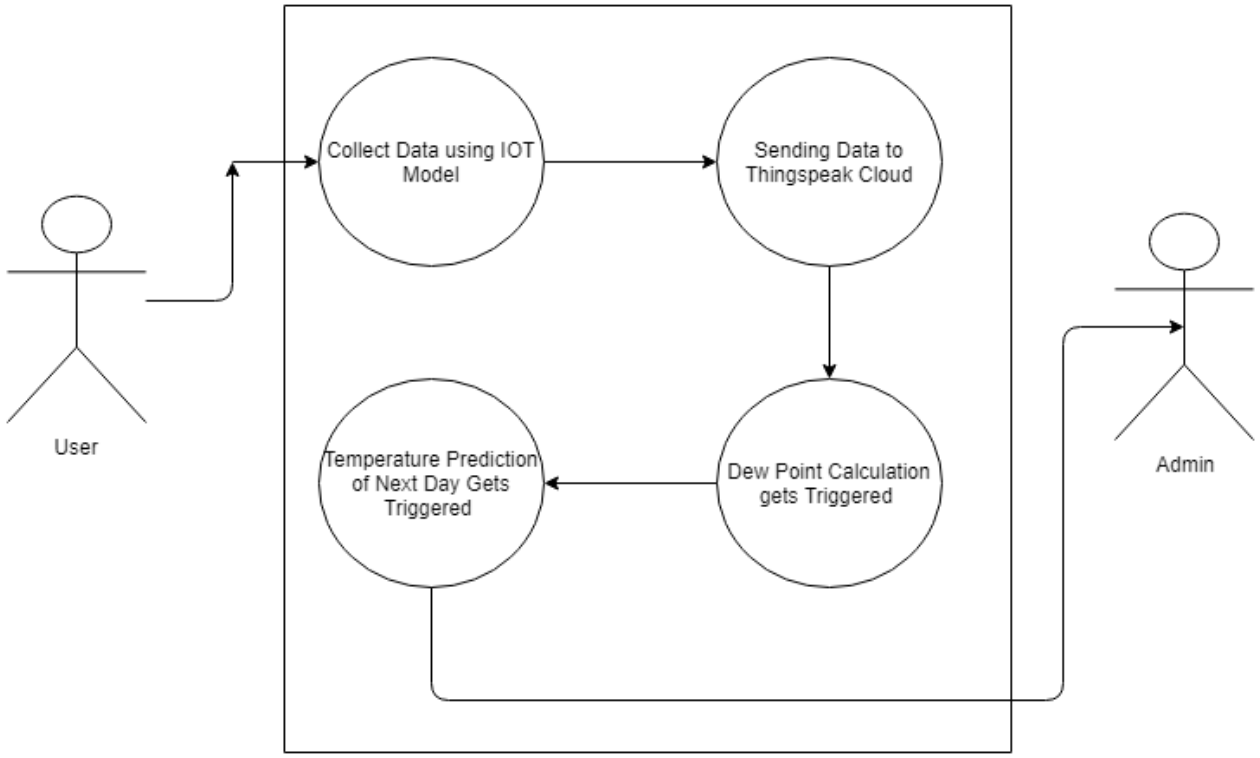
TEST2 | Arduino 1.8.9

File Edit Sketch Tools Help

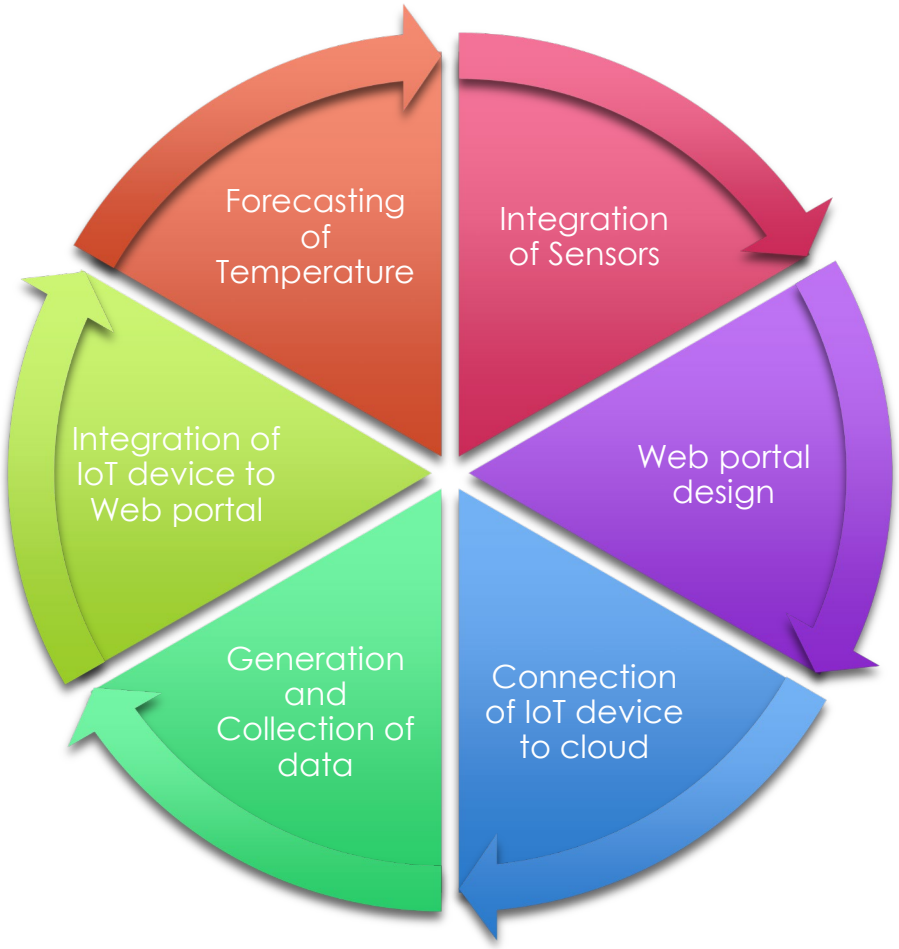
TEST2

```
#include <SFE_BMP180.h>
#include <LiquidCrystal.h>
#include <LiquidCrystal_I2C.h>
#include <SoftwareSerial.h>
#include "dht.h"


#define DHT11_An A0 // Analog Pin A0 connected to DHT11
#define DEBUG true
#define RX 10
#define TX 11

String API = "C709W7195XWIAEMO";
String HOST = "api.thingspeak.com";
String PORT = "80";

LiquidCrystal_I2C lcd(0x27, 2, 1, 0, 4, 5, 6, 7, 3, POSITIVE);  // Set the LCD I2C address, if it's not working try 0x27.
SoftwareSerial esp8266(RX,TX);

String AP = "Soham";
String PASS = "barcelona";

float MQ135_val=0 , MQ7_val=0 , DHT11_Temp=0 , DHT11_Humi=0 , Dew_val=0 , LDR_val=0 , BMP180_T=0 , BMP180_P=0 , BMP180_comp=0 ;
const int MQ135_An=1, MQ7_An=2 , LDR_An=3 , Rain_An=4;
int Altitude = 312 , i = 0 ,  Rain_val=0;  //current altitude in meters
double T=0, P=0;
int countTrueCommand;
int countTimeCommand;
boolean found = false;
```

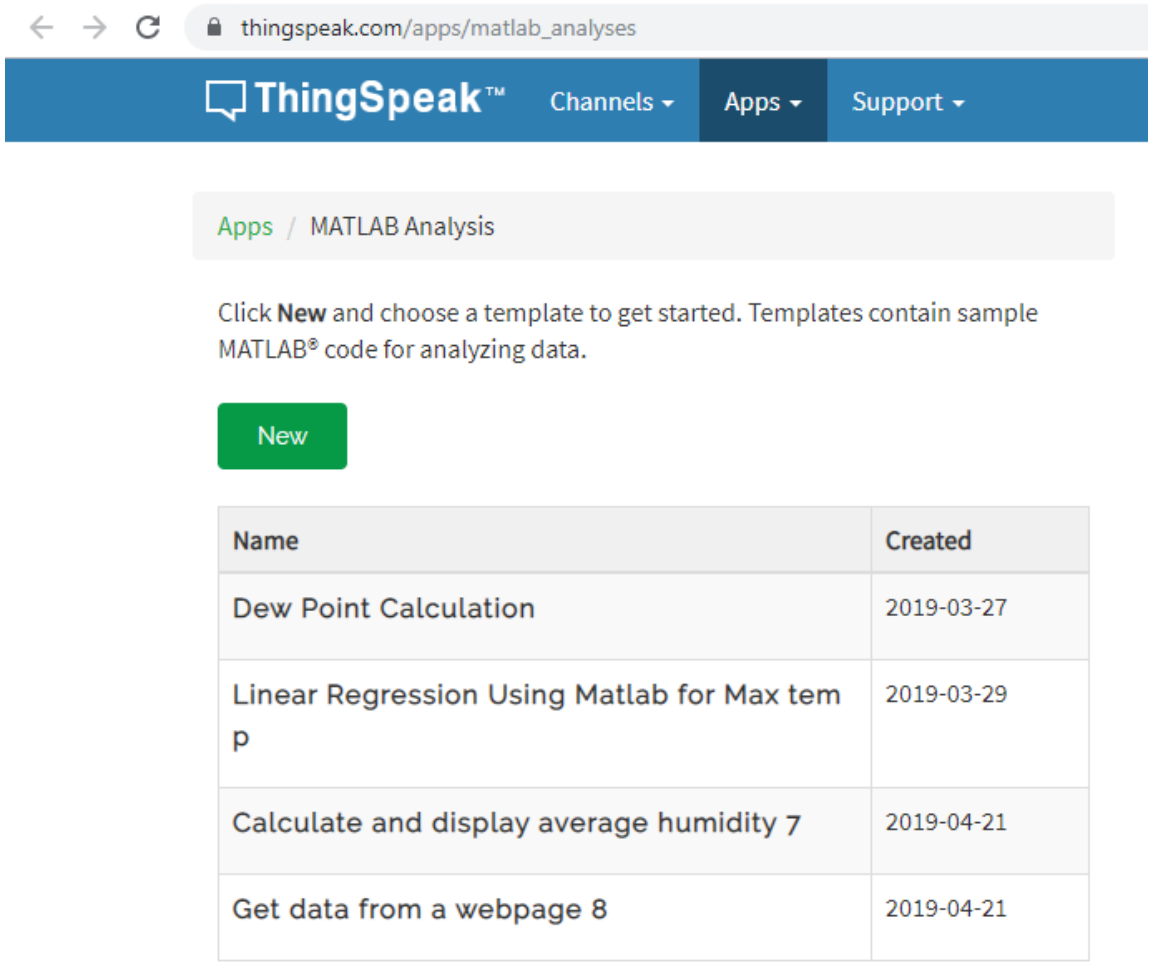
Use Case Diagram



Work done


Forecasting of Temperature
Integration of Sensors
Web portal design
Connection of IoT device to cloud
Generation and Collection of data
Integration of IoT device to Web portal

Work done


thingspeak.com/apps/matlab_analyses
ThingSpeak™
Channels
Apps
Support
Apps / MATLAB Analysis
Click New and choose a template to get started. Templates contain sample MATLAB® code for analyzing data.
New
Name
Created
Dew Point Calculation
2019-03-27
Linear Regression Using Matlab for Max tem p
2019-03-29
Calculate and display average humidity 7
2019-04-21
Get data from a webpage 8
2019-04-21

Dataset Collection

- Dataset was collected from weather underground which is currently not providing free service. The dataset provides the data from 1st May 2016 to 11th March 2018. It gives a list of parameters:

- Meantemp
- Meandewpoint
- Meanpressure
- Maxhumidity
- Minhumidity
- Maxtemp
- Mintemp
- Maxdewpoint
- Mindewpoint
- Maxpressure
- Minpressure

date	meantem	maxtempi	mintempr	meantem	meantem	meantem	meandew	meandew	meandew	meanpres	meanpres	meanpres	maxhumic	maxhumic	maxhumic	minhumic	minhumic	minhumic	maxtempi
5/4/2016	34	41	27	35	36	34	6	4	-1	1006	1005.46	1005.63	27	21	24	5	6	4	41
5/5/2016	31	38	24	34	35	36	7	6	4	1005.65	1006	1005.46	29	27	21	6	5	6	41
5/6/2016	28	34	21	31	34	35	11	7	6	1007.94	1005.65	1006	61	29	27	13	6	5	38
5/7/2016	30	38	23	28	31	34	13	11	7	1008.39	1007.94	1005.65	69	61	29	18	13	6	34
5/8/2016	34	41	26	30	28	31	10	13	11	1007.62	1008.39	1007.94	50	69	61	8	18	13	38
5/9/2016	34	42	27	34	30	28	8	10	13	1006.73	1007.62	1008.39	32	50	69	7	8	18	41
########	34	41	27	34	34	30	11	8	10	1005.75	1006.73	1007.62	45	32	50	7	7	8	42
########	32	40	25	34	34	34	16	11	8	1007.1	1005.75	1006.73	51	45	32	12	7	7	41
########	34	42	27	32	34	34	16	16	11	1006.78	1007.1	1005.75	66	51	45	16	12	7	40
########	34	42	26	34	32	34	13	16	16	1003.83	1006.78	1007.1	58	66	51	9	16	12	42
########	35	42	28	34	34	32	11	13	16	1003.56	1003.83	1006.78	40	58	66	9	9	16	42
########	36	44	29	35	34	34	13	11	13	1004.74	1003.56	1003.83	48	40	58	8	9	9	42
########	36	43	28	36	35	34	12	13	11	1005.02	1004.74	1003.56	40	48	40	7	8	9	44
########	35	44	26	36	36	35	13	12	13	1003.68	1005.02	1004.74	51	40	48	6	7	8	43
########	36	45	27	35	36	36	13	13	12	1001.58	1003.68	1005.02	42	51	40	8	6	7	44
########	38	46	29	36	35	36	11	13	13	999.88	1001.58	1003.68	45	42	51	5	8	6	45
########	38	46	30	38	36	35	6	11	13	998.96	999.88	1001.58	33	45	42	4	5	8	46
########	38	45	32	38	38	36	10	6	11	997.92	998.96	999.88	38	33	45	4	4	5	46
########	38	44	31	38	38	38	12	10	6	998.41	997.92	998.96	30	38	33	6	4	4	45
########	34	42	25	38	38	38	16	12	10	999.88	998.41	997.92	46	30	38	10	6	4	44
########	32	40	23	34	38	38	18	16	12	1001.13	999.88	998.41	69	46	30	18	10	6	42
########	34	40	28	32	34	38	16	18	16	1002.15	1001.13	999.88	74	69	46	14	18	10	40
########	34	40	28	34	32	34	17	16	18	1000.02	1002.15	1001.13	51	74	69	16	14	18	40
########	36	42	30	34	34	32	16	17	16	1001.7	1000.02	1002.15	46	51	74	14	16	14	40
########	35	40	30	36	34	34	13	16	17	1003.43	1001.7	1000.02	40	46	51	7	14	16	42
########	33	41	25	35	36	34	16	13	16	1003.17	1003.43	1001.7	46	40	46	15	7	14	40
########	32	37	26	33	35	36	16	16	13	1002.18	1003.17	1003.43	48	46	40	16	15	7	41
########	34	41	28	32	33	35	17	16	16	1003.67	1002.18	1003.17	58	48	46	21	16	15	37
6/1/2016	35	42	28	34	32	33	16	17	16	1005.08	1003.67	1002.18	52	58	48	12	21	16	41
6/2/2016	36	42	30	35	34	32	13	16	17	1005.44	1005.08	1003.67	42	52	58	10	12	21	42
6/3/2016	38	43	32	36	35	34	13	13	16	1004.8	1005.44	1005.08	38	42	52	10	10	12	42
6/4/2016	38	44	31	38	36	35	13	13	13	1004.93	1004.8	1005.44	34	38	42	10	10	10	43
6/5/2016	38	45	31	38	38	36	5	13	13	1004.67	1004.93	1004.8	27	34	38	4	10	10	44
6/6/2016	36	44	27	38	38	38	6	5	13	1003.88	1004.67	1004.93	27	27	34	4	4	10	45
6/7/2016	38	45	30	36	38	38	13	6	5	1003.29	1003.88	1004.67	51	27	27	5	4	4	44
6/8/2016	38	44	31	38	36	38	12	13	6	1002.83	1003.29	1003.88	47	51	27	4	5	4	45
6/9/2016	36	44	29	38	38	36	12	12	13	1001.24	1002.83	1003.29	43	47	51	7	4	5	44
########	37	43	31	36	38	38	12	12	12	1000.57	1001.24	1002.83	34	43	47	6	7	4	44
########	34	41	27	37	36	38	15	12	12	1000.6	1000.57	1001.24	41	34	43	11	6	7	43
########	34	41	27	34	37	36	18	15	12	1001.62	1000.6	1000.57	62	41	34	16	11	6	41
########	35	40	30	34	34	37	18	18	15	1001.2	1001.62	1000.6	58	62	41	23	16	11	41
########	34	39	29	35	34	34	19	18	18	1001.42	1001.2	1001.62	58	58	62	17	23	16	40
########	34	39	30	34	35	34	18	19	18	1000.63	1001.42	1001.2	49	58	58	18	17	23	39
########	35	40	30	34	34	35	18	18	19	998.76	1000.63	1001.42	52	49	58	16	18	17	39
########	36	42	31	35	34	34	17	18	18	999.57	998.76	1000.63	46	52	49	21	16	18	40
########	36	42	30	36	35	34	17	17	18	1000.55	999.57	998.76	41	46	52	15	21	16	42
########	35	41	29	36	36	35	19	17	17	1001.7	1000.55	999.57	46	41	46	15	15	21	42
########	35	41	29	35	36	36	20	19	17	1002.2	1001.7	1000.55	52	46	41	17	15	15	41
########	34	40	27	35	35	36	21	20	19	1001.2	1002.2	1001.7	66	52	46	21	17	15	41
########	32	37	27	34	35	35	22	21	20	1001.45	1001.2	1002.2	89	66	52	27	21	17	40
########	34	40	27	32	34	35	23	22	21	1003	1001.45	1001.2	84	89	66	31	27	21	37
########	34	40	28	34	32	34	23	23	22	1002.88	1003	1001.45	79	84	89	27	31	27	40
########	34	40	29	34	34	32	21	23	23	1001.19	1002.88	1003	62	79	84	24	27	31	40
########	34	39	29	34	34	34	21	21	23	1002.21	1001.19	1002.88	62	62	79	24	24	27	40
########	34	39	29	34	34	34	21	21	21	1001.93	1002.21	1001.19	66	62	62	30	24	24	39
########	30	35	25	34	34	34	23	21	21	1001.27	1001.93	1002.21	84	66	62	35	30	24	39
########	32	39	26	30	34	34	25	23	21	1002.41	1001.27	1001.93	94	84	66	45	35	30	35
########	36	42	30	32	30	34	24	25	23	1002.48	1002.41	1001.27	84	94	84	32	45	35	39
7/1/2016	28	33	24	36	32	30	23	24	25	1000.22	1002.48	1002.41	62	84	94	22	32	45	42
7/2/2016	26	31	22	28	36	32	25	23	24	999.43	1000.22	1002.48	100	62	84	55	22	32	33
7/3/2016	28	31	24	26	28	36	25	25	23	999.87	999.43	1000.22	100	100	62	66	55	22	31

Dataset

Dataset

date meantempm maxtempm mintempm
meantempm_1 meantempm_2 meantempm_3
meandewptm_1 meandewptm_2 meandewptm_3
meanpressurem_1 meanpressurem_2 meanpressurem_3
maxhumidity_1 maxhumidity_2 maxhumidity_3
minhumidity_1 minhumidity_2 minhumidity_3
maxtempm_1 maxtempm_2 maxtempm_3 mintempm_1
mintempm_2 mintempm_3 maxdewptm_1
maxdewptm_2 maxdewptm_3 mindewptm_1
mindewptm_2 mindewptm_3 maxpressurem_1
maxpressurem_2 maxpressurem_3 minpressurem_1
minpressurem_2 minpressurem_3 precipm_1
precipm_2 precipm_3

5/4/2016 34 41 27 35 36 34 6 4 -1 1006
1005.46 1005.63 27 21 24 5 6 4 41 43
43 29 29 26 12 10 9 -2 -2 -10 1009
1008 1009 1000 1001 999 0 0 0

5/5/2016 31 38 24 34 35 36 7 6 4 1005.65
1006 1005.46 29 27 21 6 5 6 41 41
43 27 29 29 13 12 10 0 -2 -2 1008
1009 1008 1001 1000 1001 0 0
0

Feature Selection

- For finding out the relevant parameters we conducted various test:

- Correlation Check: Correlation is a statistical technique that can show whether and how strongly pairs of variables are related. For example, height and weight are related; taller people tend to be heavier than shorter people.
- P Value Check: The p-value is the level of marginal significance within a statistical hypothesis test representing the probability of the occurrence of a given event. The p-value is used as an alternative to rejection points to provide the smallest level of significance at which null hypothesis would be rejected. A smaller p-value means that there is stronger evidence in favor of the alternative hypothesis.

1. **Backward Stepwise Regression** : It is a stepwise regression approach, that begins with a full (saturated) model and at each step gradually eliminates variables from the regression model to find a reduced model that best explains the data. Also known as Backward Elimination regression.

2. **R-Squared and Adj R-Squared:** R-squared measures the proportion of the variation in your dependent variable (Y) explained by your independent variables (X) for a linear regression model. Adjusted R-squared adjusts the statistic based on the number of independent variables in the model.

- After conducting the following analysis the following parameters were identified:

1. For max temperature prediction:

   Maxtempm_1 , Maxpressure_1 , Mintempm_3 , Maxpressure_3 , Meanpressurem_3

2. For min temperature prediction:

   Mintempm_1 , Meantempm_3 , Maxtempm_1 ,Maxtempm_2 ,Maxdewptm_1 ,Maxdewptm_3 ,Meandewptm_1 ,Meandewptm_2 ,Meanpressure_1

3. For mean temperature prediction:

   Meantemp_3 , Meanpressure_3 , Maxtemp_1 , Maxtemp_2 ,Mintemp_1

Results

Parameters	Current Temperature Accuracy is 99.05%		
	Next day Min Temperature Accuracy	Next day Max Temperature Accuracy	Next day Mean Temperature Accuracy
Explained variance	0.94	0.93	0.95
Mean absolute error	1.35 degree Celsius	1.29 degree Celsius	1.10 degree Celsius
Median absolute error	1.09 degree Celsius	0.97 degree Celsius	0.90 degree Celsius

Results


Jaipur,Rajasthan MAX
32
Jaipur,Rajasthan Tomorrow MAX
32
Jaipur,Rajasthan MIN
26
Jaipur,Rajasthan Tomorrow MIN
23
Activate Windows
Go to Settings to activate Windows.

Result Analysis

- Current Temperature Accuracy: Temperature shown contains an error of 0.5 degree Celsius.

- Next day Min Temperature Accuracy:

  Explained variance= 0.94

  Mean absolute error= 1.35 degree Celsius

  Median absolute error= 1.09 degree Celsius

- Next day Max Temperature Accuracy:

  Explained variance= 0.93

  Mean absolute error= 1.29 degree Celsius

  Median absolute error= 0.97 degree Celsius

- Next day Mean Temperature Accuracy:

  Explained variance= 0.95

  Mean absolute error= 1.10 degree Celsius

  Median absolute error= 0.90 degree Celsius

Future Scope

- The model is able to sense the values and display current day's temperature on the web portal. As of now, the model is also able to predict next day's min and max temperature.
- For the future, the model is expected to forecast weather and predict temperature for next few days. The model is also expected to calculate and predict air quality.
- This model can further be extended to be used to study about the working conditions in different work environments and how they were impacting the work efficiency of the employees.

Conclusion

- The real-time data i.e. time-series data is gathered, cleansing and analysis is performed on this dataset using Data Analytics.
- Here the values are predicted by implementing linear regression using MATLAB, the prediction can be done for next few days. The proposed system to train the model takes dataset of last two years.
- Finally, after cleansing of data, training the model, it finally displays current and predicted values of the temperature on the web portal.

Have a nice WEATHER!!!

Thank you