SITE DOWN PREDICTION USING MACHINE LEARNING ALGORITHM

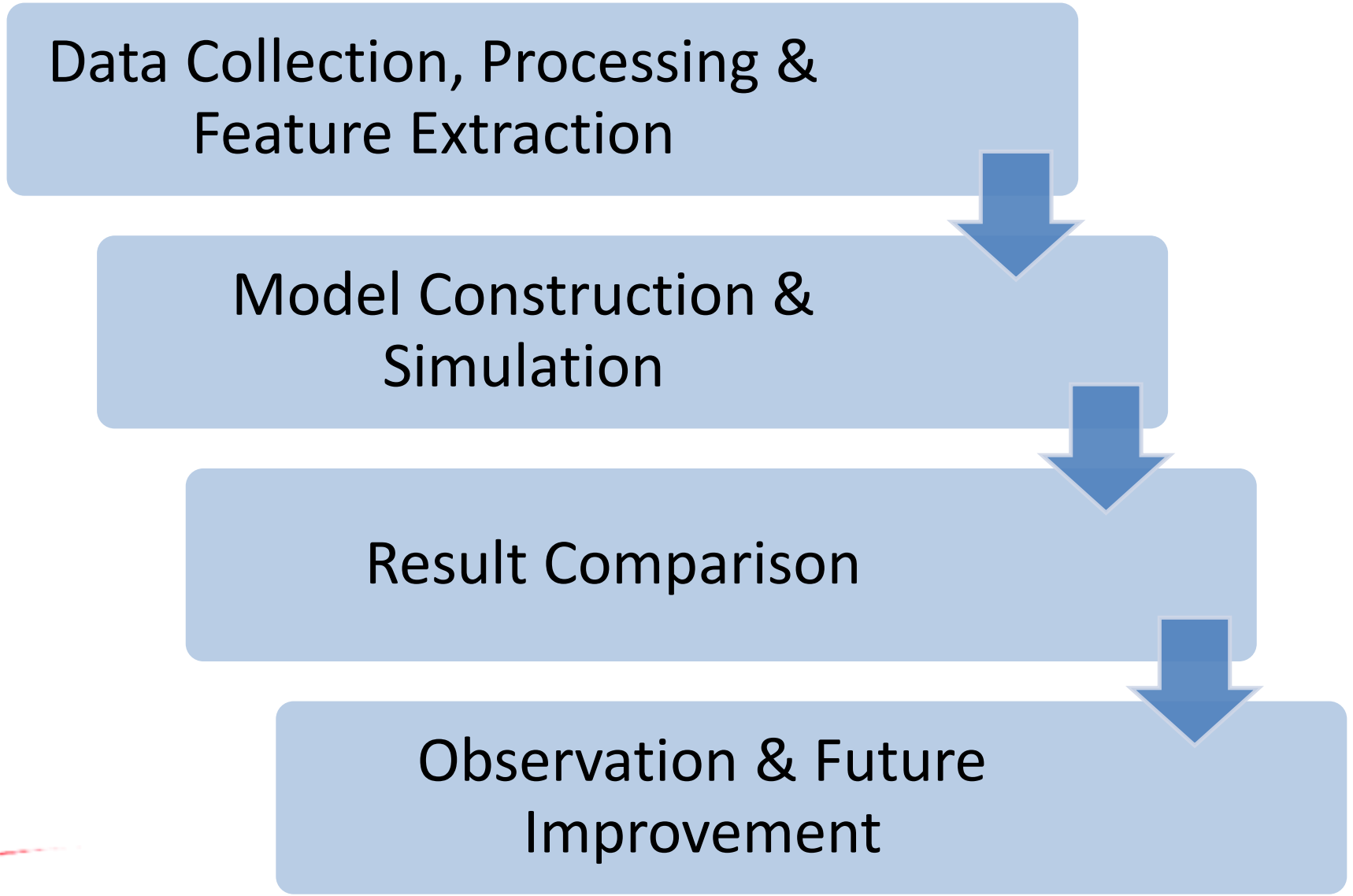
PREPARED BY:

MD MAHMUD FERDOUS

TECHNOLOGY

ROBI AXIATA LIMITED

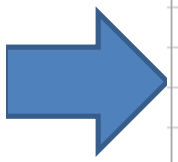
PROCESS FLOW



DATA COLLECTION AND PROCESSING

Update of Site Down at 17/06/2019 09:00:00 AM

Region:: 2G/3G/4G/P1
National:: 65/62/50/13
DHK_M: 1/2/9/0
DHK_N: 4/5/1/0
DHK_S: 0/2/1/0
CTG_M: 1/2/7/0
CTG_N: 5/5/2/1
CTG_S: 4/5/6/2
COM: 0/1/0/0
NOA: 2/1/0/0
SYL: 12/11/0/1
MYM: 7/8/0/0
BAR: 1/1/1/1
KHL: 1/0/0/0
KUS: 1/2/4/1
RAJ: 3/4/5/1
RANG: 23/13/14/6



Date	Time	2G	3G	4G	P1	X	Data_4G	Data_3G	Data_2G	Data_P1
16/6/2019	21:00:00	21	45	37	9	21	21,37	21,45	21,21	21,9
17/6/2019	0:00:00	26	33	44	6	0	0,44	0,33	0,26	0,6
17/6/2019	6:00:00	37	47	36	11	6	6,36	6,47	6,37	6,11
17/6/2019	9:00:00	65	62	50	13	9	9,50	9,62	9,65	9,13
17/6/2019	12:00:00	46	54	40	6	12	12,40	12,54	12,46	12,6
17/6/2019	15:00:00	48	49	34	6	15	15,34	15,49	15,48	15,6
17/6/2019	18:00:00	42	55	32	8	18	18,32	18,55	18,42	18,8
17/6/2019	21:00:00	54	59	33	9	21	21,33	21,59	21,54	21,9
18/6/2019	6:00:00	98	101	56	31	6	6,56	6,101	6,98	6,31
18/6/2019	9:00:00	99	140	60	24	9	9,60	9,140	9,99	9,24
18/6/2019	12:00:00	64	74	49	15	12	12,49	12,74	12,64	12,15
18/6/2019	15:00:00	41	60	43	8	15	15,43	15,60	15,41	15,8
18/6/2019	18:00:00	30	53	44	6	18	18,44	18,53	18,30	18,6
18/6/2019	21:00:00	54	42	41	13	21	21,41	21,42	21,54	21,13
19/6/2019	0:00:00	37	47	43	9	0	0,43	0,47	0,37	0,9
19/6/2019	6:00:00	242	92	61	126	6	6,61	6,92	6,242	6,126

MODEL CONSTRUCTION & SIMULATION

ML Algorithm

Linear Regression

Gradient Descent

The objective of linear regression is to minimize the cost function

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

where the hypothesis $h_\theta(x)$ is given by the linear model

$$h_\theta(x) = \theta^T x = \theta_0 + \theta_1 x_1$$

$$\theta_j := \theta_j - \alpha\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)} \quad \text{(simultaneously update } \theta_j \text{ for all } j\text{)}.$$

Programming Environment	Feature	Parameters
• GNU Octave (open source)	• Single (Time Series) which is very popular and powerful in predictive analysis (i.e. weather forecasting)	• Training Set: **25** • Test Set: **5** • Iteration: **1500** • Learning Rate (α): **0.01**

C:\Users\Noman\Desktop\Ferdous\ML\Octave\OCTAVE~1.0\mingw64\bin\octave-gui...

```
Plotting Data ...
Program paused. Press enter to continue.

Testing the cost function ...
With theta = [0 ; 0]
Cost computed = 387.920000

With theta = [-1 ; 2]
Cost computed = 445.220000
Program paused. Press enter to continue.

Running Gradient Descent ...
Theta found by gradient descent:
23.351945
-0.707106
Program paused. Press enter to continue.
Visualizing J(theta_0, theta_1) ...
octave:18>
```

Figure 1

File Edit Tools

Z+ Z- Insert Text Axes Grid Autoscale

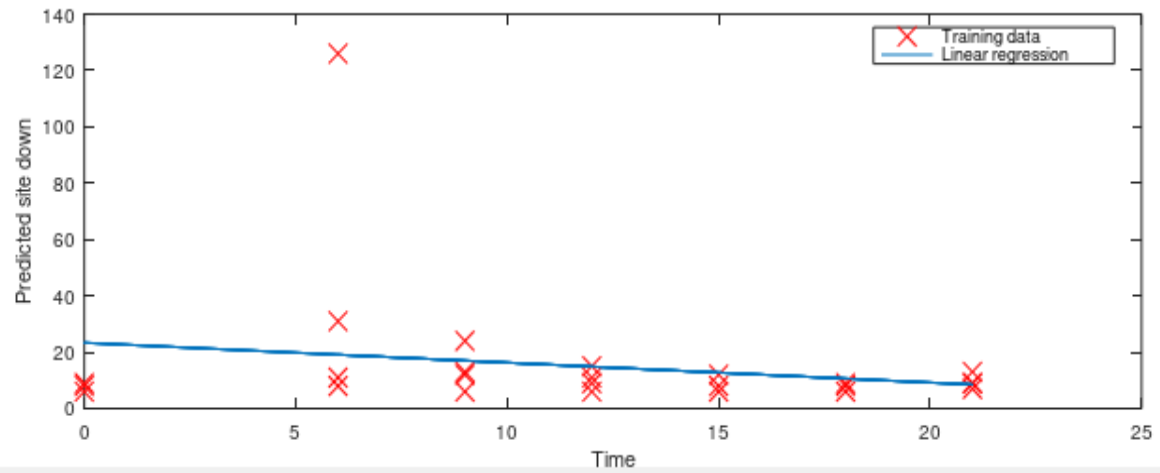
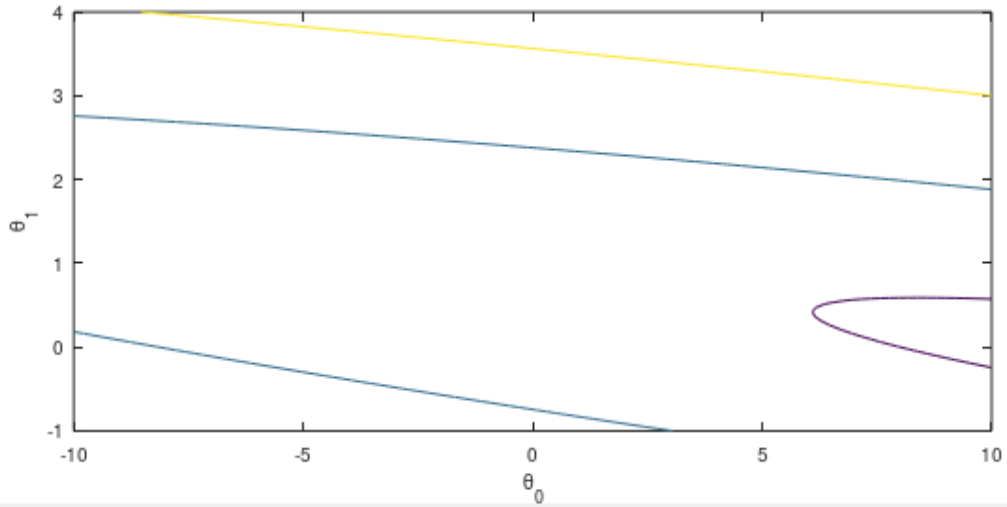


Figure 3

File Edit Tools

Z+ Z- Insert Text Axes Grid Autoscale

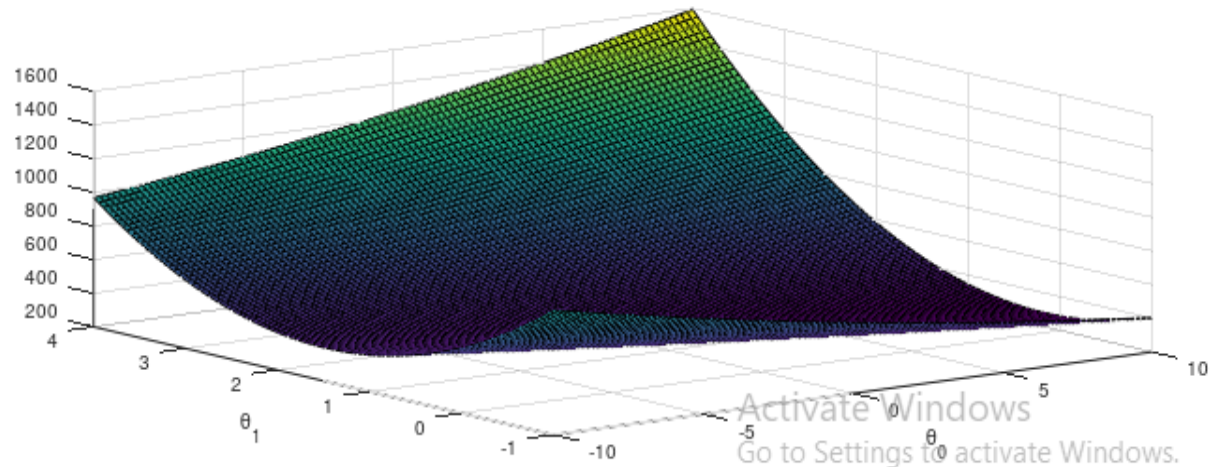


(-8.5774, 3.7895)

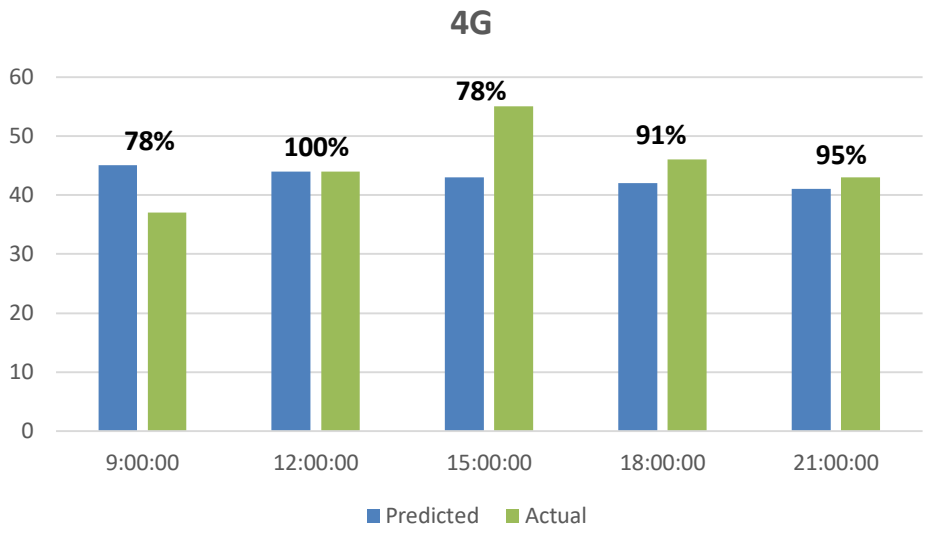
Figure 2

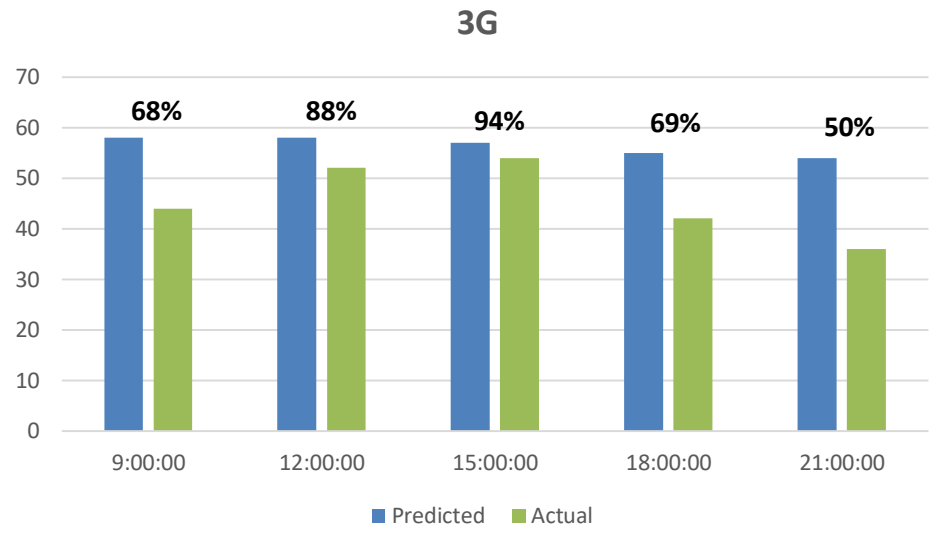
File Edit Tools

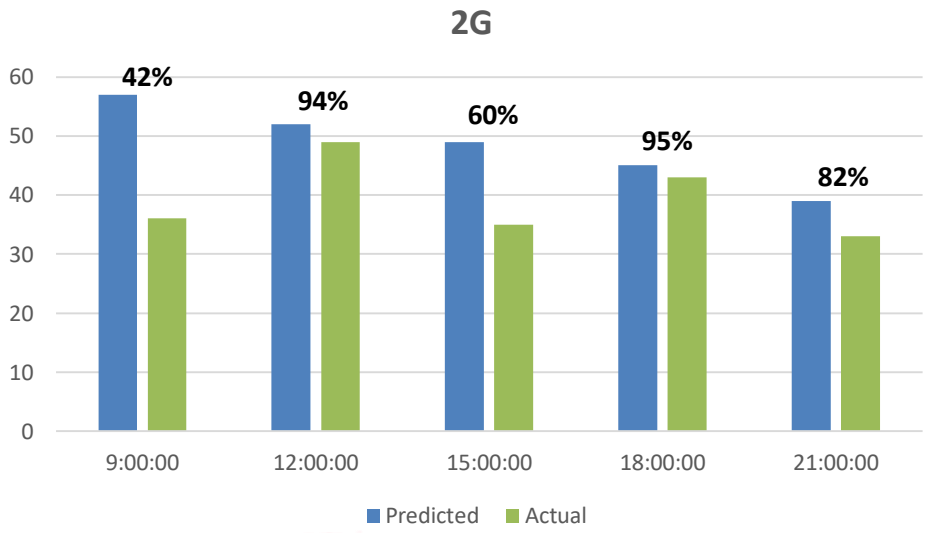
Z+ Z- Insert Text Axes Grid Autoscale

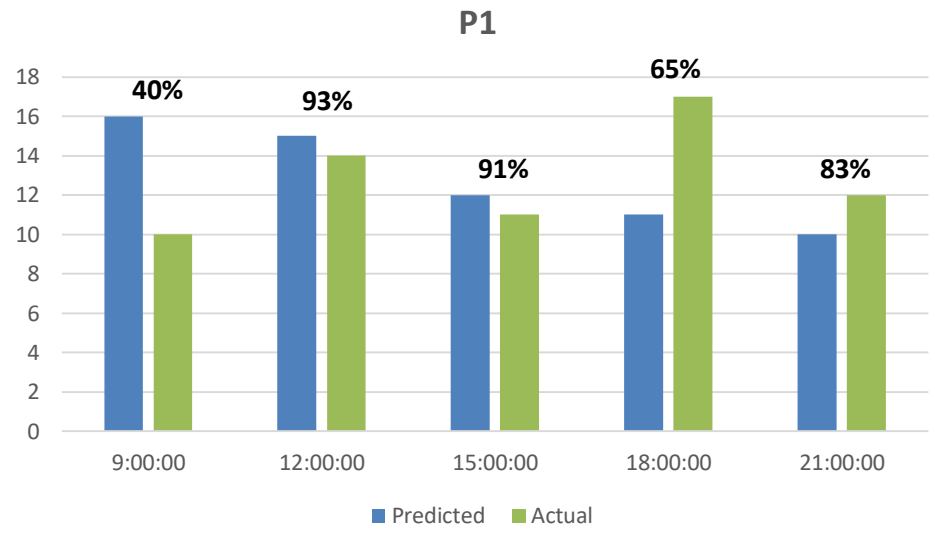


RESULT COMPARISON



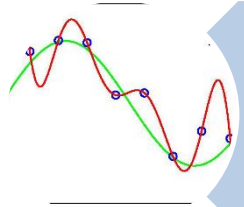






Observation & Future Improvement

Currently external data to feed in the model is messy which need to be stored in a structured way. It is the most important & biggest challenge right now.



Under fitting problem occurred for certain time stamp due to low training sample which can be optimized by feeding more sample as well as non-linear analysis.

For now, prediction is only for national total. Have a plan to breakdown region wise also.

In this algorithm single feature used. In future will add more features (i.e. Mains Fail, DC Low, Weather: Temperature, rain, wind) and it will predict more accurately.

THANK YOU