FINAL REVIEW

Guide - Dr. M. Jaya Bharata Reddy

- Angad Bajwa (107118014)
- Mandar Burande (107118056)
- Aditya Pethkar (107118072)
- Priyansh Joshi (107118076)

INDEX

TITLE

Cyber Attack Detection in Power System SCADA networks using Machine Learning Techniques

OBJECTIVES OF THE WORK

01

To monitor and analyze real-time data flow in SCADA networks

02

To detect and thwart various incoming cyber attacks such as man-in-the-middle and remote tripping commands, etc.

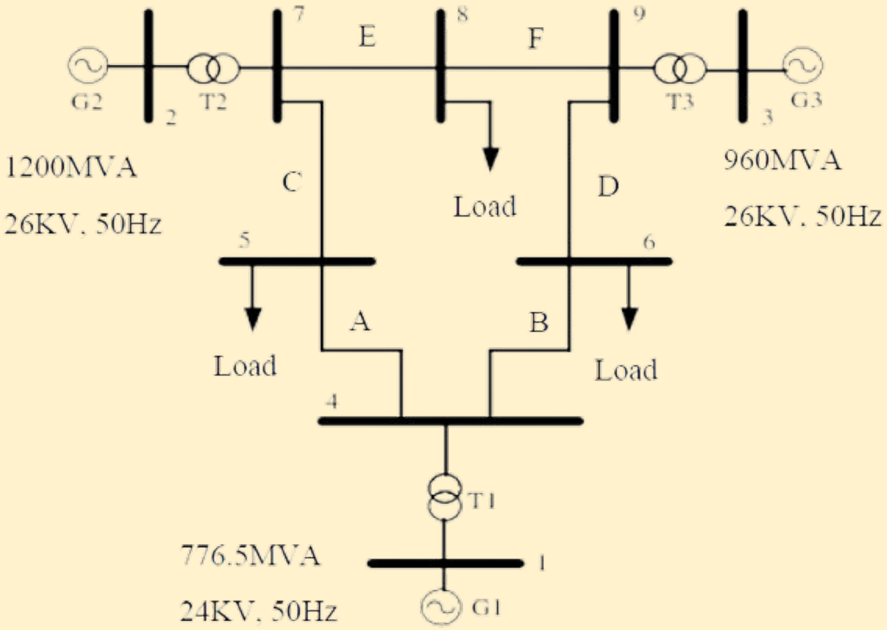
03

To put forth inferences to assist in implementing further solutions

EXISTING METHODOLOGIES

- **Fault classification using machine learning techniques -**
  - Classifier models used to segregate fault and non-fault data as well as types of fault exist.
  - Data from power system simulation is collected and and labelled.
  - Data is fed into machine learning models like K-Means Clustering.
  - This allows the model to classify and predict the fault and/or type of fault occurring.
- **Cyber attacks on SCADA networks -**
  - Existing research deals with cyber threats and attacks on SCADA networks, Modbus protocol, etc
  - Cyber attacks are simulated on targeted network topologies.
  - Vulnerabilities are exposed and reported.
  - Solutions are proposed to counter the vulnerabilities.

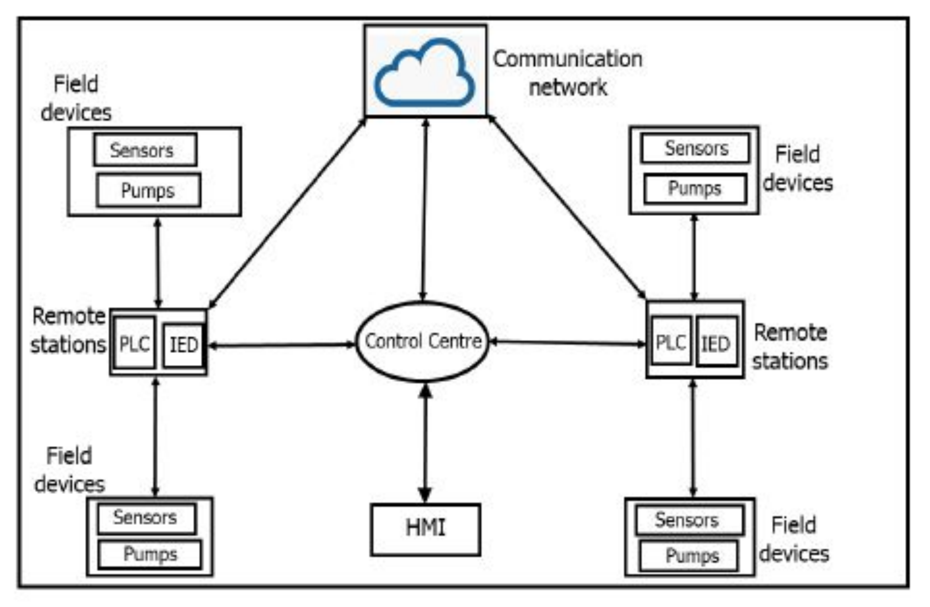
SINGLE LINE DIAGRAM



Single Line Diagram of the IEEE Standard 9-Bus System

Parameters	Ratings
Generator 1	776.5 MVA, 24 KV,50 Hz
Generator 2	1200 MVA, 26 KV, 50 Hz
Generator 3	960 MVA, 26 KV, 50 Hz
Line A	150 Km
Line B	120 Km
Line C	120 Km
Line D	140 Km
Line E	110 Km
Line F	110 Km

SCADA ARCHITECTURE

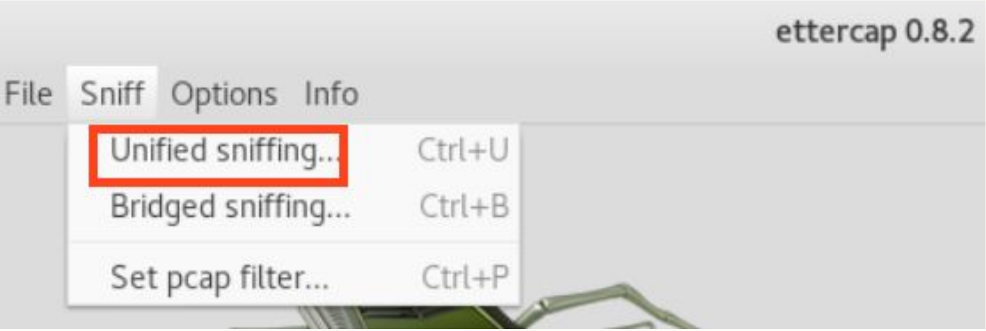


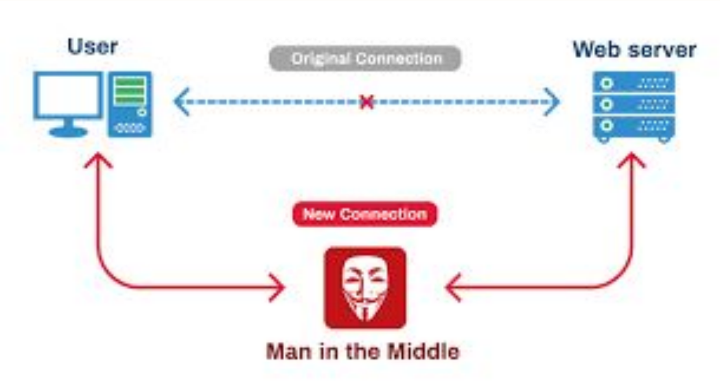
SCADA Architecture

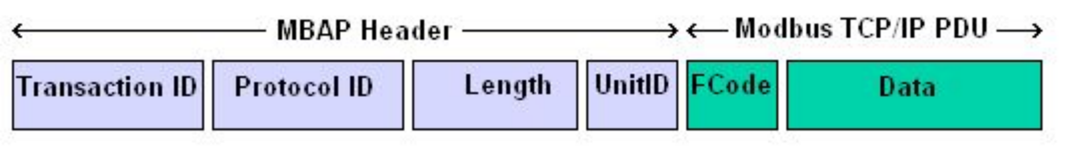
METHODOLOGY-I

Fault Location	LG	LLG	LL	LLL	LLLG
a	20-130,40-110,60-90 (in Km)	20-130,40-110,60-90 (in Km)	20-130,40-110,60-90 (in Km)	20-130,40-110,60-90 (in Km)	20-130,40-110,60-90 (in Km)
b	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)
c	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)	20-100,40-80,60-60 (in Km)
d	20-120,40-100,60-80 (in Km)	20-120,40-100,60-80 (in Km)	20-120,40-100,60-80 (in Km)	20-120,40-100,60-80 (in Km)	20-120,40-100,60-80 (in Km)
e	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)
f	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)	20-90,40-70,60-50 (in Km)

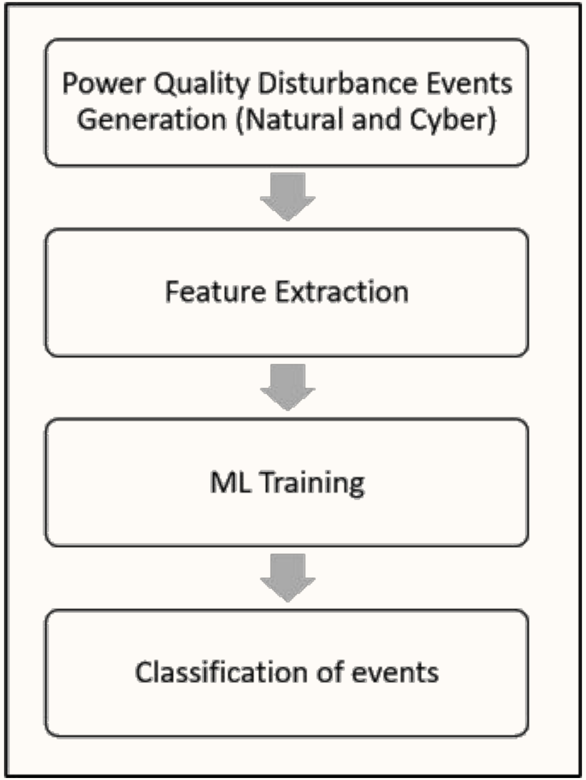
METHODOLOGY-II







METHODOLOGY-III



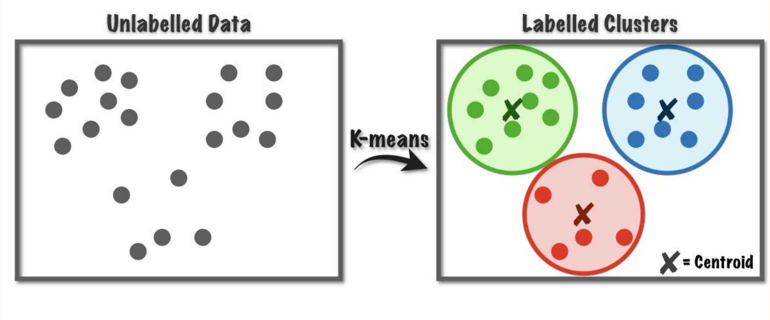
$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN}$$

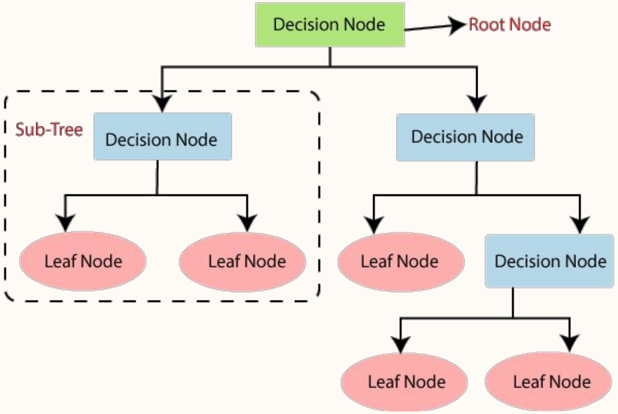
$$Precision = \frac{TP}{TP + FP}$$

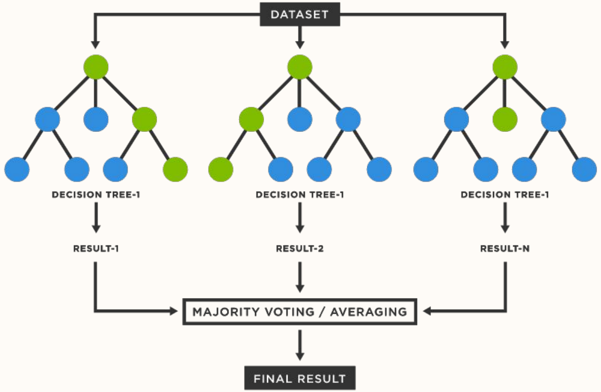
$$Recall = \frac{TP}{TP + FN}$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

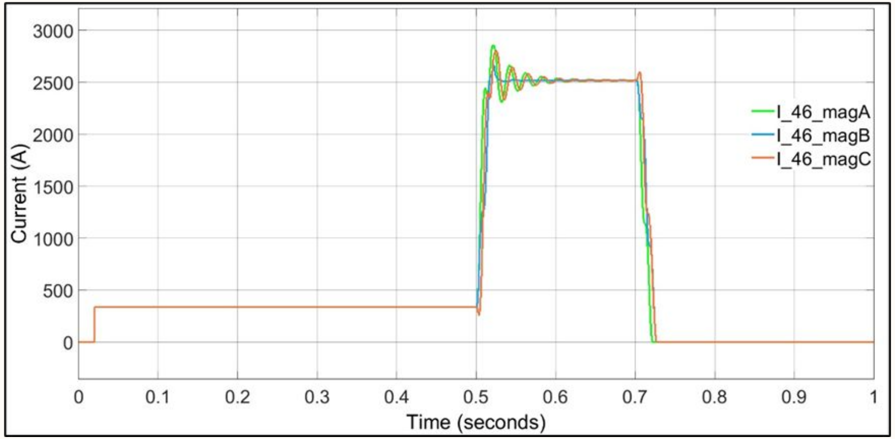
METHODOLOGY-IV



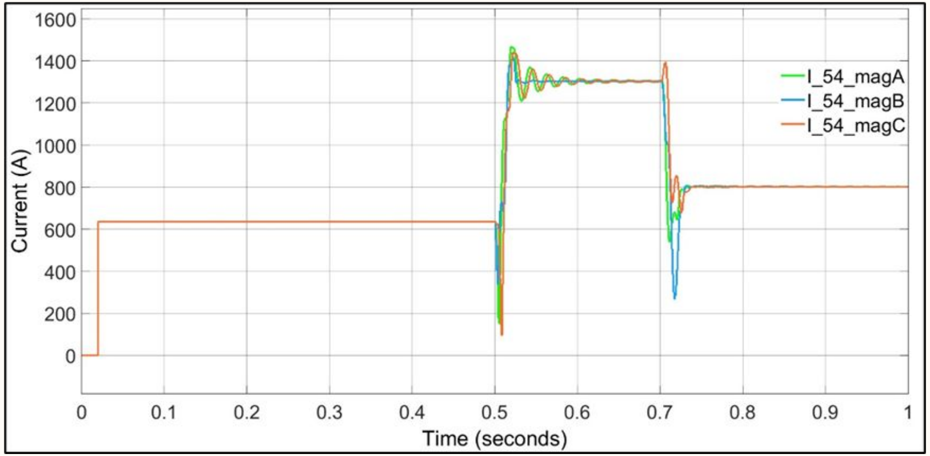




RESULTS-I



Line Containing LLL Fault (Line 4-6, B)



Line Without Fault (Line 5-4, A)

RESULTS-II

	E	F	G	H	I	J	K	L	M	N
1	V_magA	V_magB	V_magC	I_angleA	I_angleB	I_angleC	I_magA	I_magB	I_magC	Condition
2	0	0	0	0	0	0	0	0	0	0
3	339129.1274	339174.2646	339120.4549	42.32377724	-77.67682107	162.3177591	673.3249848	673.4025995	673.357707	0
4	339109.0626	339109.2964	339109.6045	42.31852025	-77.6811343	162.3179937	673.3296228	673.3391181	673.3378873	0
5	339109.3018	339109.0065	339109.3242	42.31833781	-77.68144481	162.3182603	673.330996	673.333528	673.3344754	0
6	339109.3693	339109.3052	339109.3813	42.31822754	-77.68168937	162.3182982	673.3319044	673.3315148	673.3325562	0
7	339109.3945	339109.42	339109.4322	42.31819114	-77.68181242	162.3182564	673.332436	673.331532	673.3319488	0
8	339109.4008	339109.4389	339109.4463	42.31818592	-77.68185389	162.3182292	673.3326731	673.3318221	673.3318434	0
9	339109.4029	339109.4316	339109.444	42.31818894	-77.68185876	162.3182216	673.3327207	673.33196	673.3318559	0
10	339109.4057	339109.4264	339109.4398	42.31819329	-77.68185197	162.3182219	673.3326857	673.3319965	673.3318815	0
11	339109.4089	339109.4255	339109.4373	42.31819777	-77.68184263	162.3182235	673.3326272	673.3320104	673.3319086	0
12	339109.4118	339109.4261	339109.4359	42.31820214	-77.68183327	162.3182247	673.3325673	673.3320273	673.3319379	0
13	339109.4144	339109.4268	339109.4349	42.31820621	-77.68182452	162.3182256	673.3325123	673.3320471	673.3319679	0
14	339109.4165	339109.4273	339109.434	42.31820985	-77.68181661	162.3182263	673.3324633	673.3320659	673.3319963	0
15	339109.4183	339109.4276	339109.4333	42.31821303	-77.68180963	162.3182271	673.3324204	673.3320822	673.3320217	0
16	339109.4198	339109.4279	339109.4326	42.31821577	-77.68180357	162.3182277	673.3323833	673.332096	673.3320438	0
17	339109.4211	339109.4281	339109.432	42.31821811	-77.68179836	162.3182283	673.3323515	673.3321078	673.332063	0
18	339109.4222	339109.4283	339109.4316	42.31822012	-77.6817939	162.3182288	673.3323243	673.3321179	673.3320794	0
19	339109.4232	339109.4284	339109.4312	42.31822182	-77.68179009	162.3182292	673.332301	673.3321264	673.3320934	0
20	339109.424	339109.4286	339109.4308	42.31822328	-77.68178685	162.3182295	673.3322813	673.3321337	673.3321054	0
21	339109.4247	339109.4287	339109.4305	42.31822452	-77.68178408	162.3182298	673.3322644	673.3321399	673.3321157	0
22	339109.4253	339109.4288	339109.4303	42.31822557	-77.68178172	162.3182301	673.33225	673.3321452	673.3321244	0
23	339109.4258	339109.4289	339109.4301	42.31822647	-77.68177972	162.3182303	673.3322378	673.3321498	673.3321318	0
24	339109.4262	339109.429	339109.4299	42.31822723	-77.68177801	162.3182305	673.3322274	673.3321536	673.3321382	0
25	339109.4266	339109.429	339109.4297	42.31822788	-77.68177655	162.3182307	673.3322186	673.3321569	673.3321436	0
26	339109.4269	339109.4291	339109.4296	42.31822844	-77.68177531	162.3182308	673.3322111	673.3321597	673.3321482	0
27	339109.4271	339109.4291	339109.4295	42.31822891	-77.68177426	162.3182309	673.3322047	673.332162	673.3321521	0
28	173866.5744	345305.6814	282939.1511	-26.09966641	-73.93301491	171.2511233	2165.008953	807.4377816	568.3175833	1
29	165256.8021	342706.8161	283007.2674	-27.27530904	-72.89221831	169.9198067	2040.051416	799.993996	558.4583754	1
13821	339110.3205	339110.3213	162173.9982	42.31885301	-77.68114691	162.3188531	1571.915109	4126.549346	3546.617807	2
13822	339110.3205	339110.3213	339110.3207	42.31885302	-77.68114689	162.3188531	673.3298191	673.3298246	673.3298238	2

B_75 | B_78 | B_87 | B_46 | B_64 | B_96 | B_69 | B_89 | B_98 | B_57 | B_54 | B_45

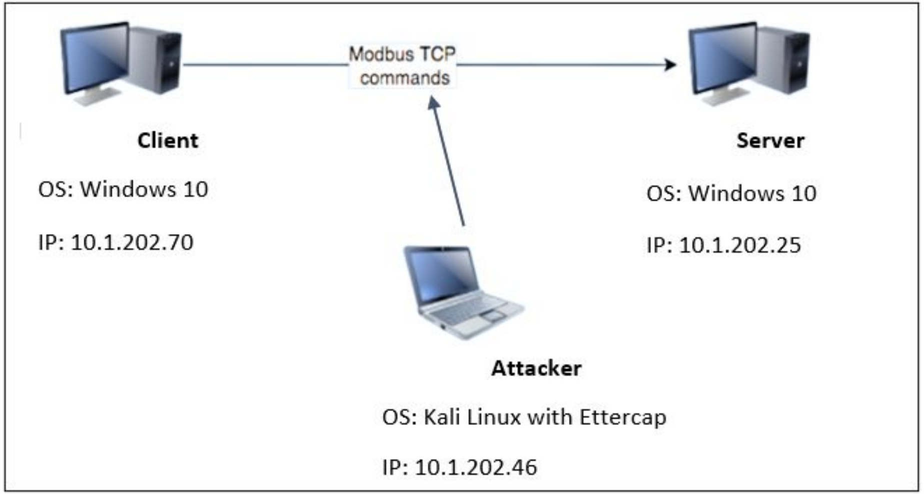
Features -

- Phase A Voltage magnitude
- Phase A Voltage Angle
- Phase B Voltage magnitude
- Phase B Voltage Angle
- Phase C Voltage magnitude
- Phase C Voltage Angle
- Phase A Current magnitude
- Phase A Current Angle
- Phase B Current magnitude
- Phase B Current Angle
- Phase C Current magnitude
- Phase C Current Angle

Conditions -

- Natural Operation (0)
- Fault (1)
- Cyber Attack (2)

RESULTS-III



Cyber Attack Architecture

```
------------------------------------------------------------
Decoded Data
------------------------------------------------------------
V_75_angleA       39.98264694213867
V_75_angleB       -80.01714324951172
V_75_angleC       159.98382568359375
V_75_magA         328342.40625
V_75_magB         328336.90625
V_75_magC         328340.09375
I_75_angleA       37.61441421508789
I_75_angleB       96.51009368896484
I_75_angleC       160.12152099609375
I_75_magA         8.639863047221752e-09
I_75_magB         3.810341375753978e-09
I_75_magC         2.925536923825689e-09
------------------------------------------------------------
------------------------------------------------------------
```

Data decoded from Modbus Payload format

RESULTS-IV

K-Means

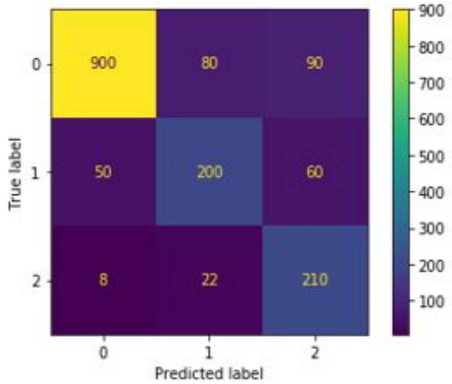
Fault Line	Accuracy	Precision	Recall	F1 Score
a	86.9	82	88	87
b	87.5	83	87	89
c	86.5	84	86	85
d	86.9	83	87	86
e	87.3	87	88	85
f	87.2	89	82	85

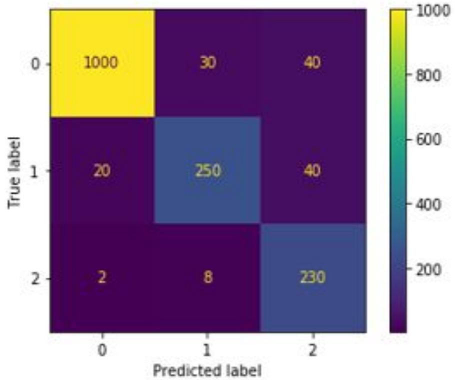
Decision Tree Classifier

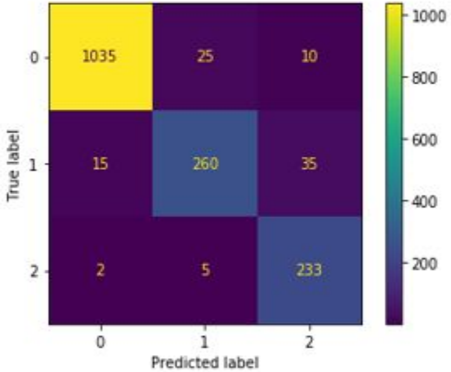
Fault Line	Accuracy	Precision	Recall	F1 Score
a	95.1	95	96	94
b	95.6	96	95	95
c	96.3	96	97	94
d	95.8	96	95	92
e	95.3	96	94	95
f	95.9	94	96	95

Random Forest

Fault Line	Accuracy	Precision	Recall	F1 Score
a	96.1	96	93	95
b	96.6	98	94	96
c	97.3	98	96	96
d	96.8	95	95	95
e	96.3	96	94	94
f	96.9	97	96	95







CONCLUSIONS - I

- Simulated multiple types of power system faults – natural as well as cyber-attacks,
- Collected relevant data from the generated data samples and
- Analysed them using multiple machine learning techniques.
- Industry-standard tools like MATLAB, Simulink, Linux, Ettercap, Wireshark, Jupyter Notebooks, sklearn, pandas, numpy and more were used during this project. This further elevates the relevance of the work done.
- Based on the observed results, inferences were gleaned and suggestions to counter as well as thwart the problems were proposed.

CONCLUSIONS - II

- Devised and proposed a methodology to simulate and investigate the occurrence as well as the impact of the different obstacles and sabotages on the system.
- One such method is to add synchronised identification fields and hash functions to the Modbus Protocol.
- This allows the protocol to be more secure and also enables authentication.
- Combined with a well-trained machine learning model, the system will become robust and capable of detecting and thwarting cyber-attacks

REFERENCES

1. **O. A. Alimi, K. Ouahada, and A. M. Abu-Mahfouz** (2020) A Review of Machine Learning Approaches to Power System Security and Stability. *IEEE Access, vol. 8, pp. 113512-113531, 2020*
2. **Lemay, Antoine, and José M. Fernandez** (2016) Providing SCADA Network Data Sets for Intrusion Detection Research. *CSET @ USENIX Security Symposium.*
3. **R. C. Borges Hink, J. M. Beaver, M. A. Buckner, T. Morris, U. Adhikari and S. Pan** (2014) Machine learning for power system disturbance and cyber-attack discrimination. *7th International Symposium on Resilient Control Systems (ISRCS)*
4. **W. Rahman, M. Ali, A. Ullah, H. Rahman, M. Iqbal, H. Ahmad, A. Zeb, Z. Ali, M. Shahzad, and B. Taj** (2012) Advancement in Wide Area Monitoring Protection and Control Using PMU's Model in MATLAB/SIMULINK. *Smart Grid and Renewable Energy, Vol. 3 No. 4*

THANK YOU!