

22MAT122 – MATEMATICS FOR COMPUTING 2

Anomaly Detection in Smart Meters

- **TEAM MEMBERS:**

- **Yashwanth B - CB.SC.U4AIE24360**
- **Rohith Ravi - CB.SC.U4AIE24350**

SMART METERS

- **Two-Way Communication**
 - Sends usage data to utility providers.
 - Receives commands or updates (like remote shut-off or firmware updates).
- **Automated Meter Reading (AMR)**
 - No need for manual readings—data is sent automatically.
 - Enables remote diagnostics and updates.
- **Data Encryption & Security**
 - Uses secure communication protocols.
 - Protects consumer data from tampering or hacking.



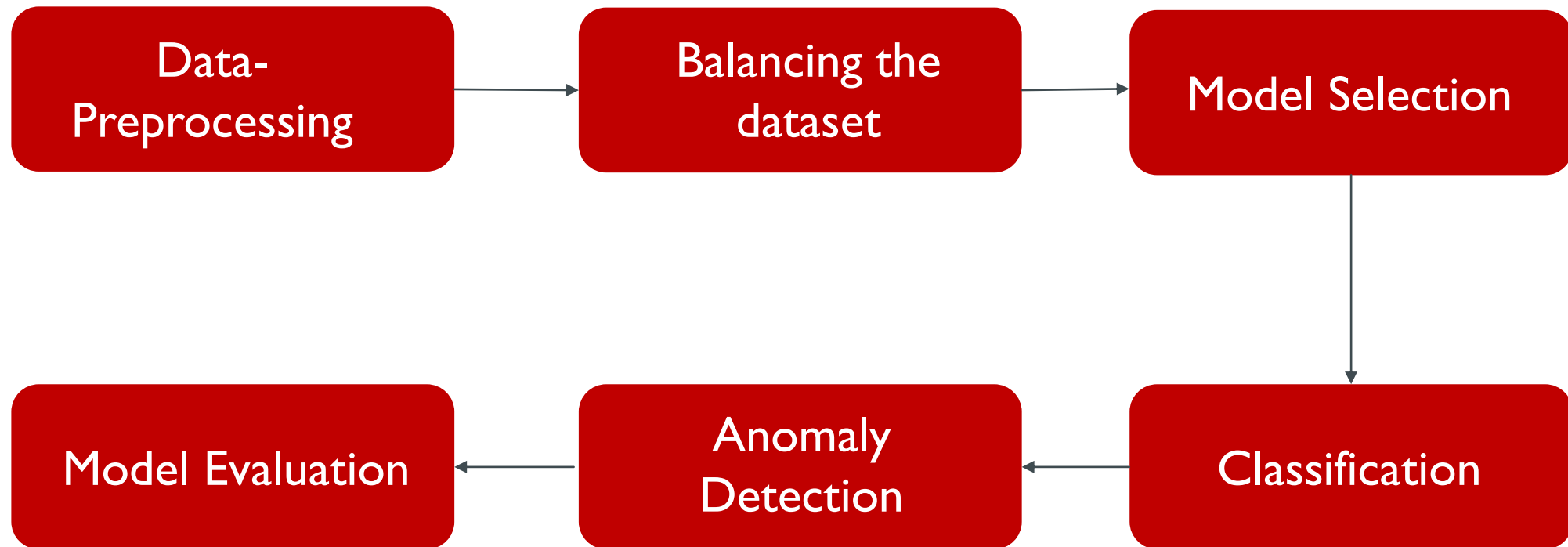
DATASET OVERVIEW

- The datasets which were used here are:
 - ❖ Normal Traffic- Contains data which does not contain any anomalies/attacks.
 - ❖ TCP_SYN attack with sniffing
 - ❖ TCP_SYN attack without sniffing
- We merged the datasets and labelled them accordingly.
- This dataset consists of 83 columns(83rd column is the label ; i.e., The Target Variable) and 219257 instances.

ABOUT THE DATASET

Category	Example Features	Purpose
Network Info	Src_ip,dst_ip,protocol	Identifies source and destination
Traffic Stats	flow_byts_s,flow_pkts_s	Measures data flow rate
Packet Details	Pkt_len_max,pkt_len_mean	Tracks packet size variations
TCP Flags	Syn_flag_cnt,ack_flag_cnt	Identifies connection behaviour
Time-Based	Flow_duration,flow_iat_mean	Detects timing irregularities
Label	Normal/Anomaly	Ground truth for classification

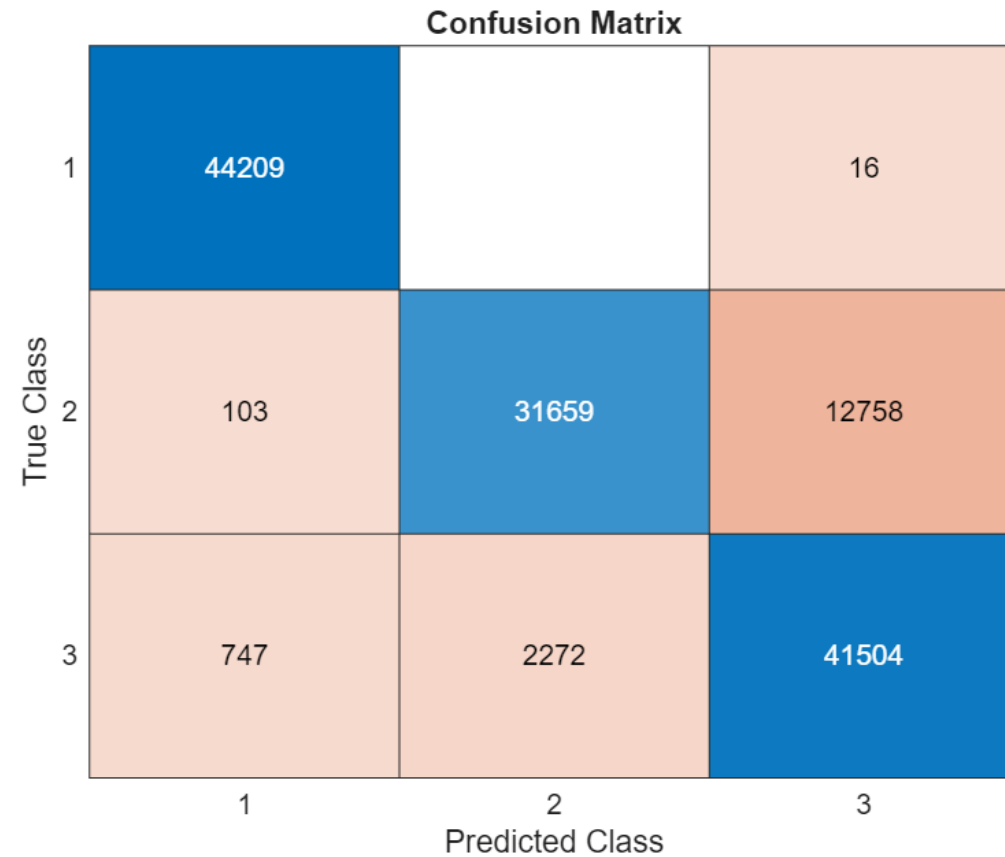
WORK FLOW



MULTICLASS CLASSIFICATION

A. MATLAB:

I. Using $Ax=b$:

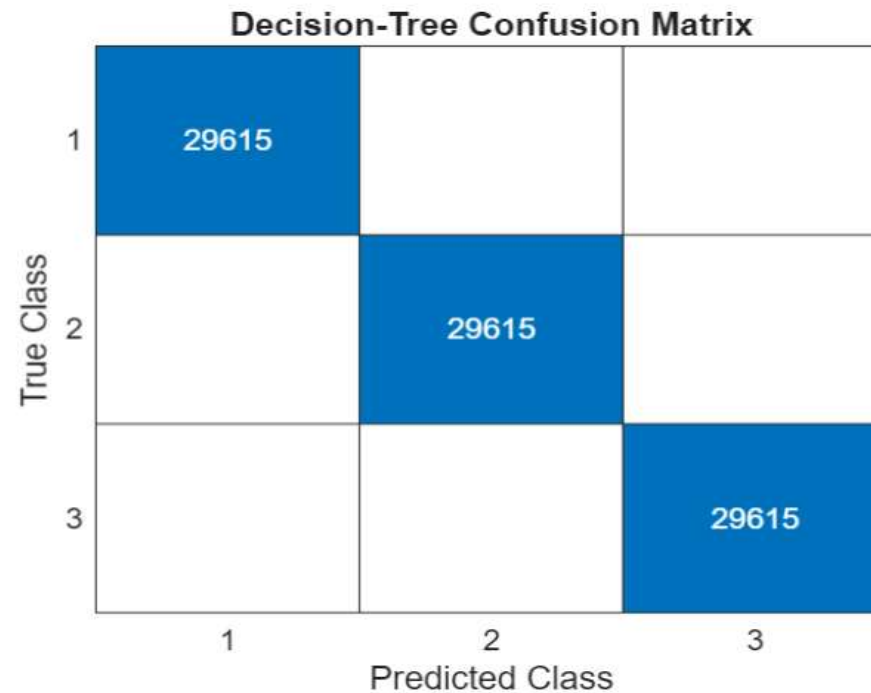


Accuracy: 88.07%

MULTICLASS CLASSIFICATION

A. MATLAB:

2.Using DecisionTree(fitctree):



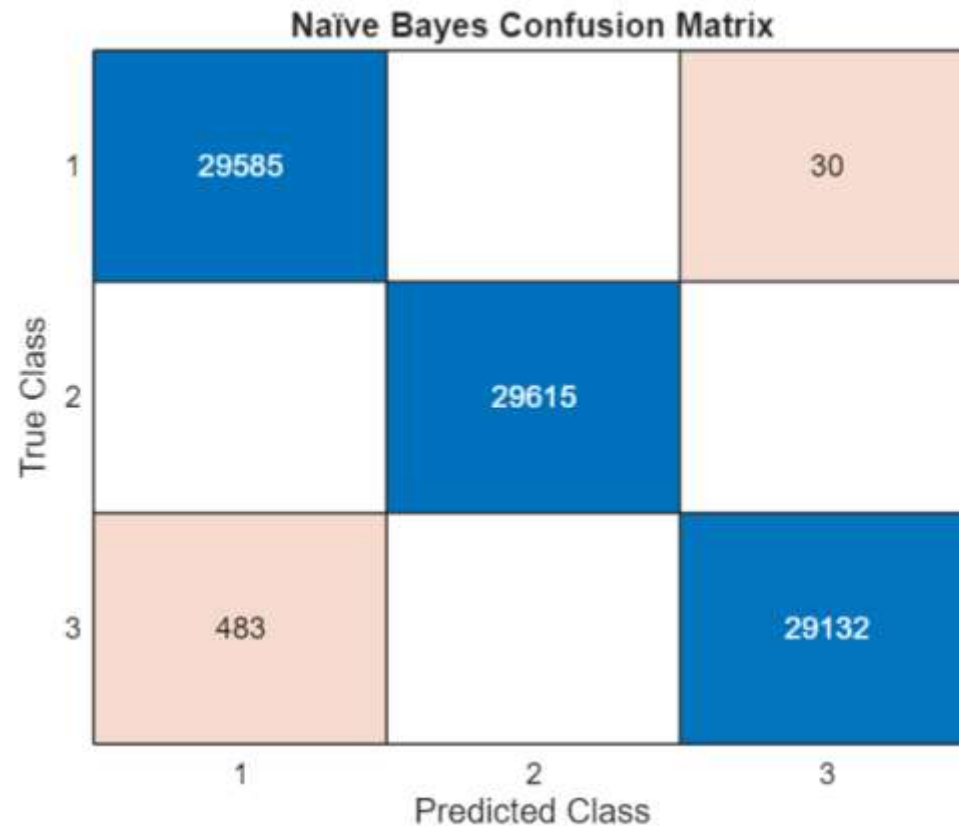
Accuracy: 100%

MULTICLASS CLASSIFICATION

A. MATLAB:

2.Using Naïve-Bayes(fitcnb):

Naïve Bayes Accuracy: 99.4226%

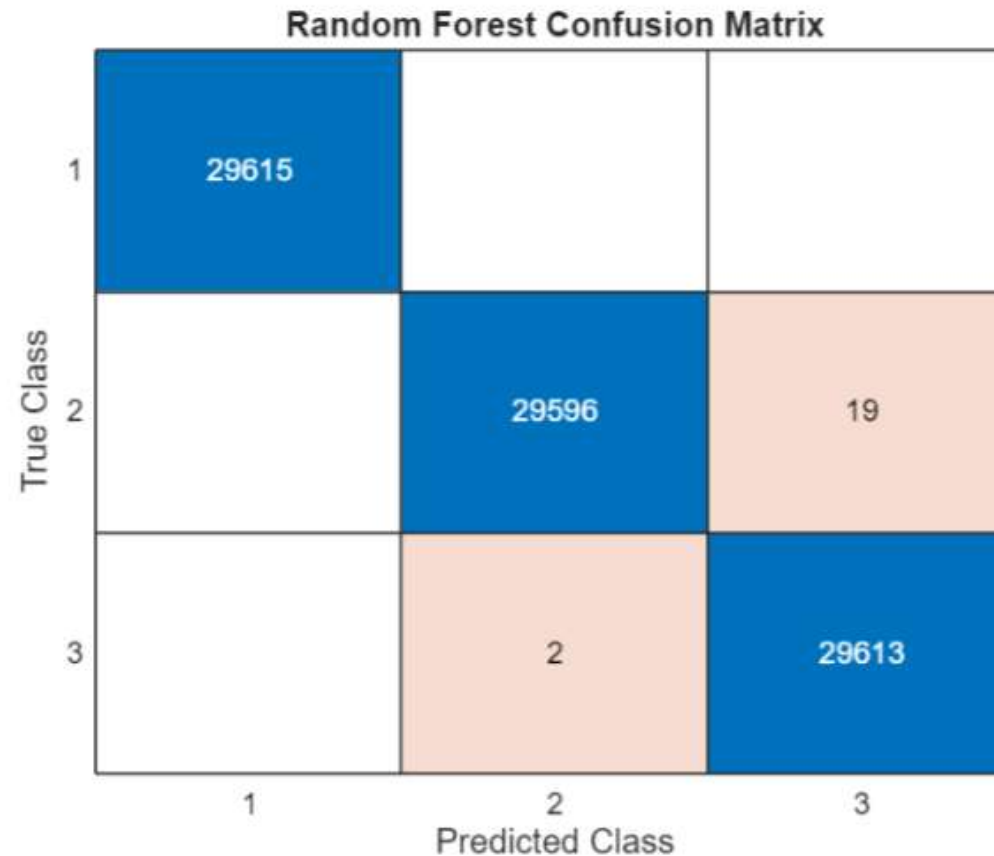


MULTICLASS CLASSIFICATION

A. MATLAB:

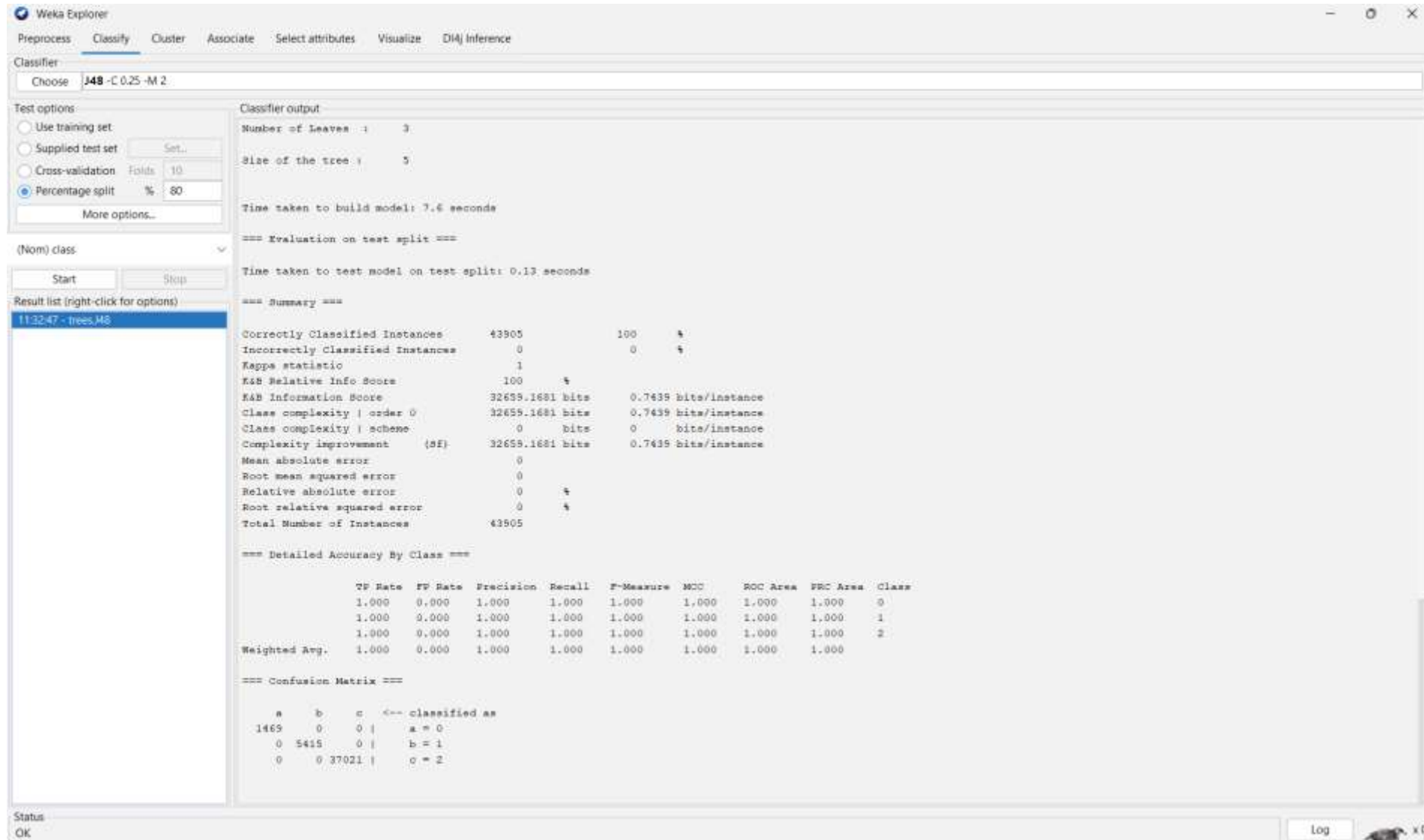
2.Using Random-Forest(Treebagger):

Random Forest Accuracy: 99.9764%



B. WEKA:

I. Using J48(Decision Tree):



The screenshot displays the Weka Explorer application window. The 'Classify' tab is active, and the 'J48 -C 0.25 -M 2' classifier is selected. The 'Test options' section shows 'Percentage split' at 80%. The 'Result list' on the left shows a single result: '11:32:47 - trees_M8'. The 'Classifier output' pane on the right displays the following information:

Classifier output

Number of Leaves : 3

Size of the tree : 5

Time taken to build model: 7.6 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.13 seconds

=== Summary ===

Metric	Value	Unit
Correctly Classified Instances	43905	
Incorrectly Classified Instances	0	
Kappa statistic	1	
K&B Relative Info Score	100	%
K&B Information Score	32655.1681	bits
Class complexity order 0	32655.1681	bits
Class complexity scheme	0	bits
Complexity improvement (SI)	32655.1681	bits
Mean absolute error	0	
Root mean squared error	0	
Relative absolute error	0	%
Root relative squared error	0	%
Total Number of Instances	43905	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	0
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	2
Weighted Avg.	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	

=== Confusion Matrix ===

a	b	c	<== classified as
1469	0	0	a = 0
0	5415	0	b = 1
0	0	37021	c = 2

Status: OK

B. WEKA:

2.Using Naïve Bayes:

The screenshot shows the Weka Explorer application window. The 'Classify' tab is selected. The 'Classifier' dropdown is set to 'NaiveBayes'. Under 'Test options', 'Percentage split' is selected with a value of 80. The 'Start' button has been clicked, and the results are displayed in the 'Classifier output' pane.

Classifier output

weight sum	7243	27178	185104
precision	8366.4962	8366.4962	8366.4962

Time taken to build model: 2.2 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 1.51 seconds

=== Summary ===

Correctly Classified Instances	43663	95.4488 %
Incorrectly Classified Instances	242	0.5512 %
Kappa statistic	0.98	
K&B Relative Info Score	97.785 %	
K&B Information Score	31935.7626 bits	0.7274 bits/instance
Class complexity order 0	32659.1681 bits	0.7439 bits/instance
Class complexity scheme	83853.3886 bits	1.9099 bits/instance
Complexity improvement (sf)	-51194.3205 bits	-1.166 bits/instance
Mean absolute error	0.0036	
Root mean squared error	0.059	
Relative absolute error	1.9663 %	
Root relative squared error	19.5756 %	
Total Number of Instances	43905	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.993	0.003	0.914	0.993	0.952	0.951	0.998	0.986	0
	0.997	0.002	0.983	0.997	0.990	0.989	0.997	0.982	1
	0.994	0.002	1.000	0.994	0.997	0.981	0.997	0.999	2
Weighted Avg.	0.994	0.002	0.995	0.994	0.995	0.981	0.997	0.997	

=== Confusion Matrix ===

a	b	c	<-- classified as
1458	0	3	a = 0
7	5398	10	b = 1
130	84	36807	c = 2

Status: OK

B. WEKA:

3.Using Random Forest:

The screenshot shows the Weka Explorer application window. The 'Classify' tab is selected. The classifier chosen is 'RandomForest' with parameters: -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1. The test options are set to 'Percentage split' at 80%. The result list on the left shows three models: 'trees.M48', 'bayes.NaiveBayes', and 'trees.RandomForest' (selected). The classifier output for 'trees.RandomForest' is displayed on the right.

Classifier output

RandomForest

Bagging with 100 iterations and base learner

`weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities`

Time taken to build model: 52.35 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.55 seconds

=== Summary ===

Correctly Classified Instances	43895	99.9863 %
Incorrectly Classified Instances	6	0.0137 %
Kappa statistic	0.9995	
K&B Relative Info Score	99.6256 %	
K&B Information Score	32536.8837 bits	0.7411 bits/instance
Class complexity order 0	32659.1681 bits	0.7439 bits/instance
Class complexity scheme	106.0561 bits	0.0024 bits/instance
Complexity improvement (BF)	32553.072 bits	0.7414 bits/instance
Mean absolute error	0.001	
Root mean squared error	0.0117	
Relative absolute error	0.5626 %	
Root relative squared error	3.8928 %	
Total Number of Instances	43905	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.996	0.000	1.000	0.996	0.998	0.998	1.000	1.000	0
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1
	1.000	0.001	1.000	1.000	1.000	0.999	1.000	1.000	2
Weighted Avg.	1.000	0.001	1.000	1.000	1.000	0.999	1.000	1.000	

=== Confusion Matrix ===

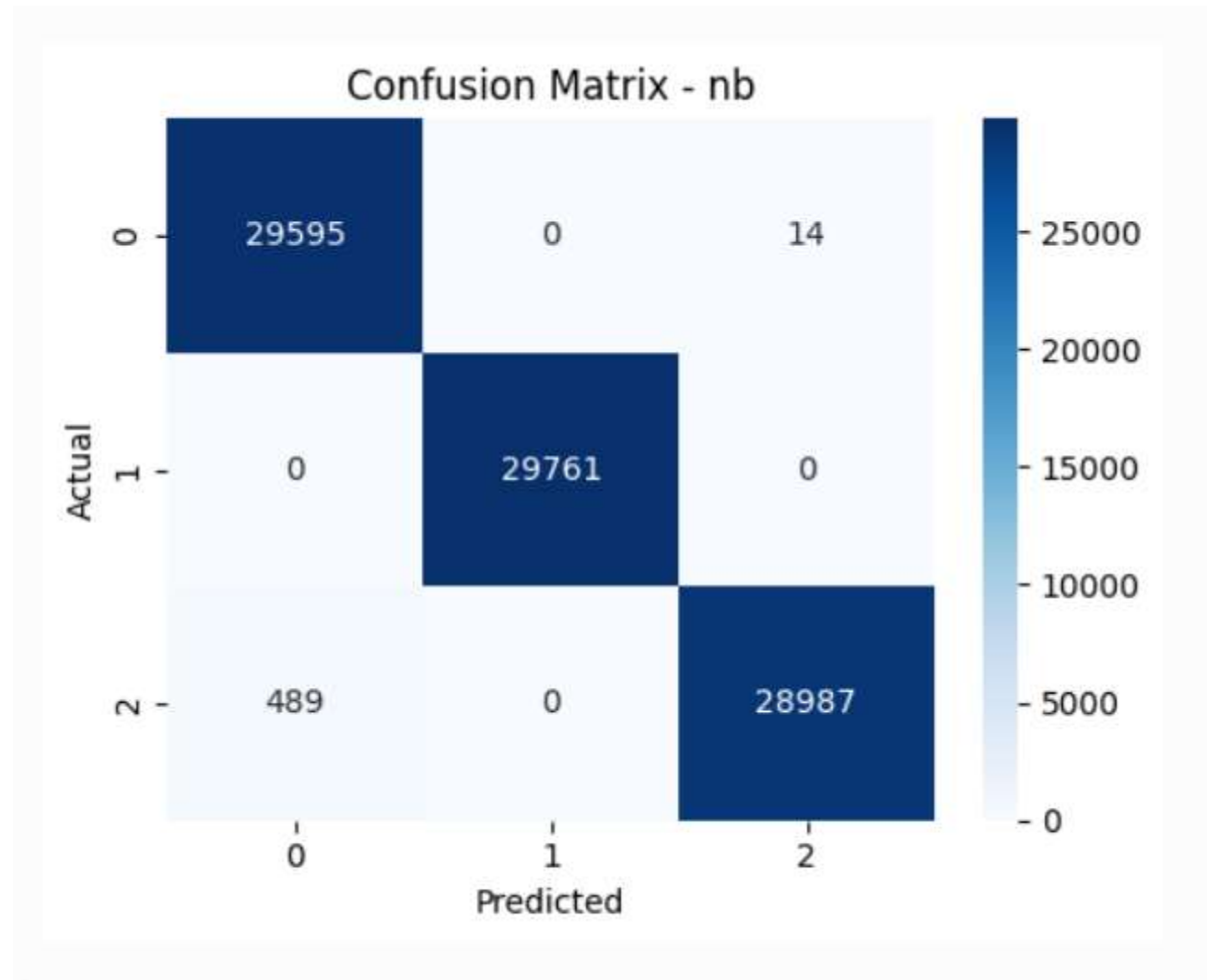
a \ b	c	<-- classified as
1463	0	a = 0
0	5415	b = 1
0	0	c = 2

Status: OK

Log

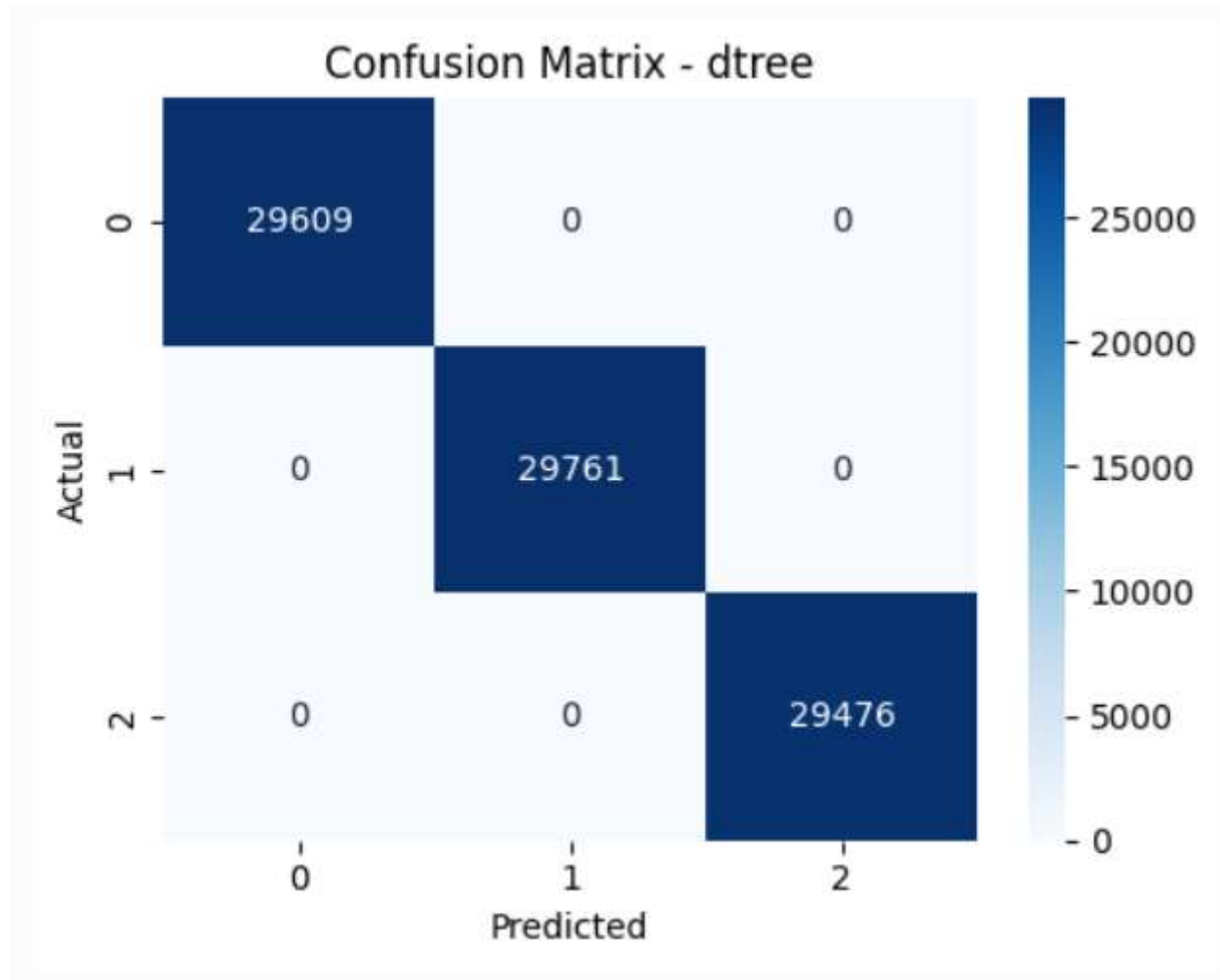
MARIMO

I.Using Naïve Bayes:



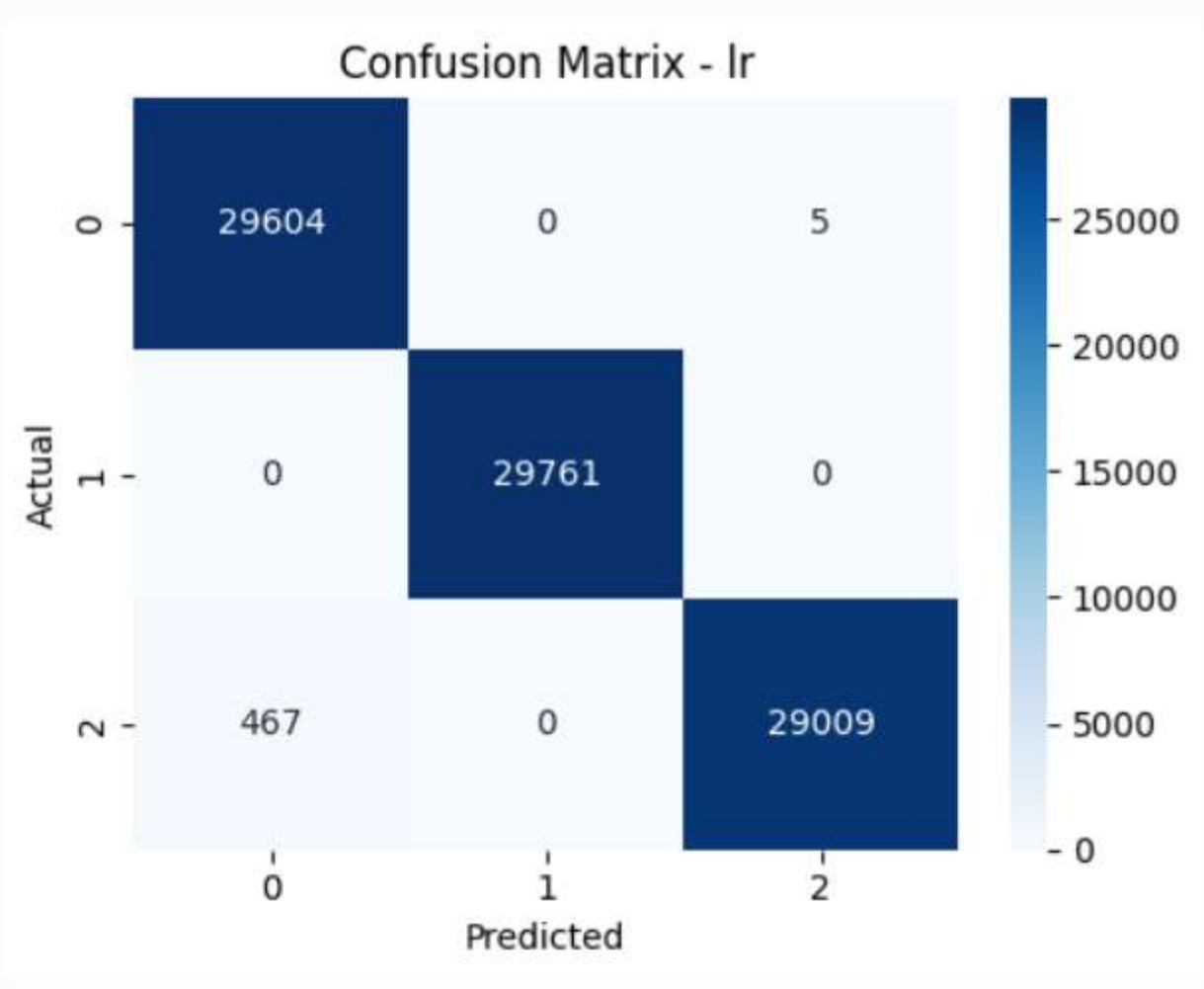
MARIMO

2.Using Decision Tree:



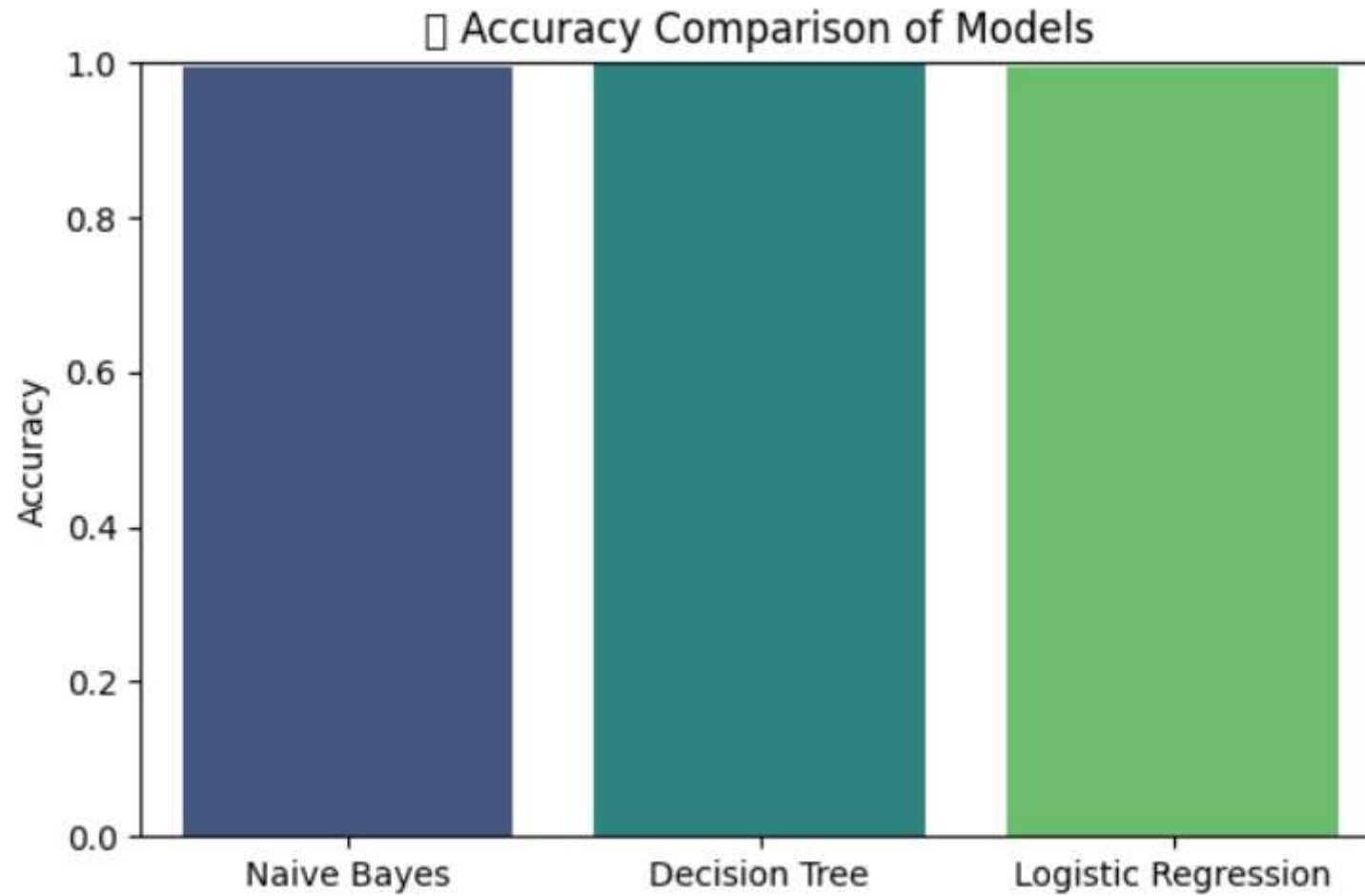
MARIMO

3.Using Logistic Regression:

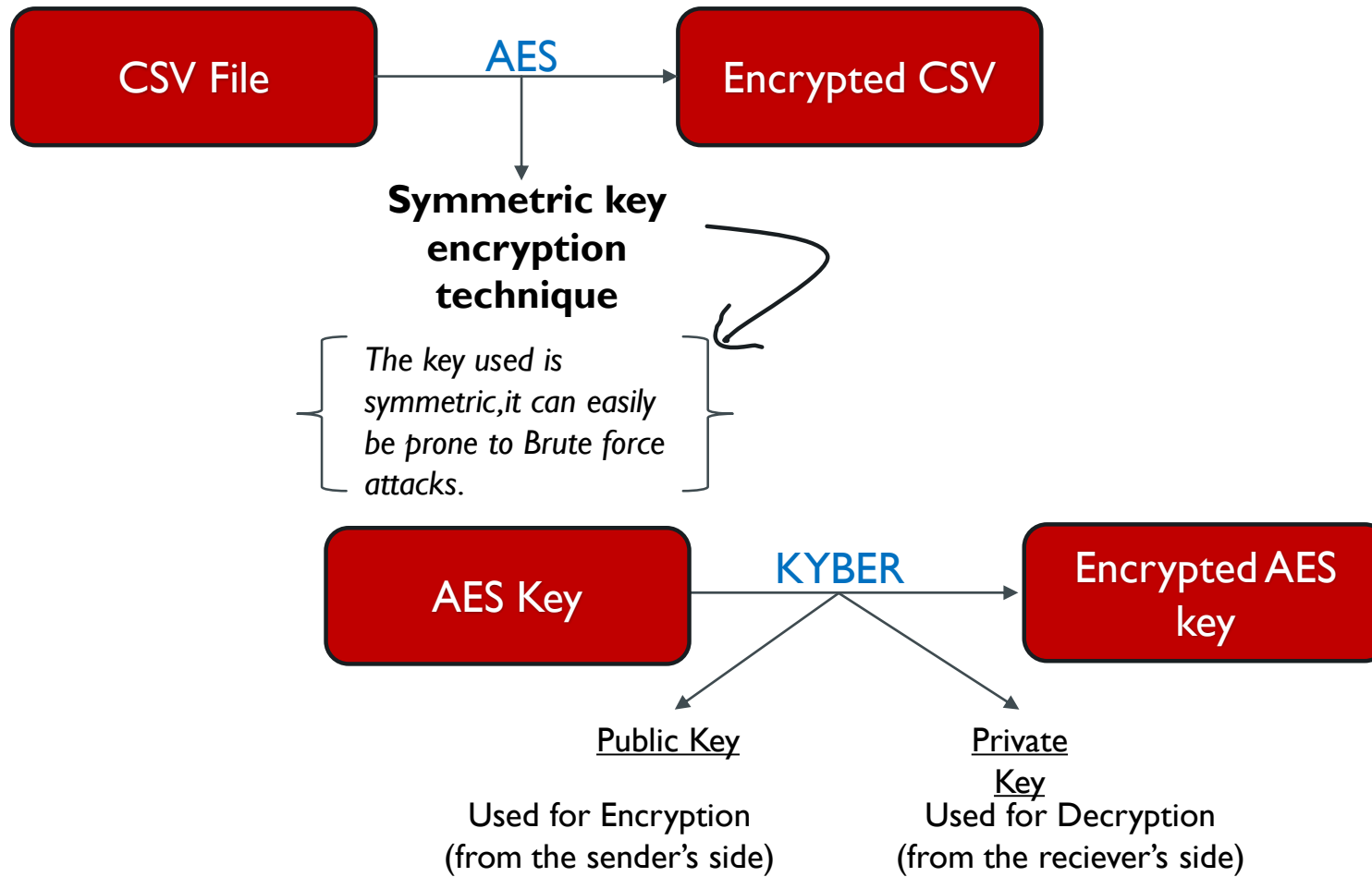


MARIMO

Model Accuracy Comparison:



ENCRYPTION



REFERENCES

- [1] M.A. Hossain and M. S. Islam, "Enhancing DDoS attack detection with hybrid feature selection and ensemble-based classifier: A promising solution for robust cybersecurity," *Measurement: Sensors*, vol. 32, p. 101037, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2665917424000138>
- [2] M. Mohurle and V.V. Panchbhai, "Review on realization of AES encryption and decryption with power and area optimization," *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, Delhi, India, 2016, pp. 1-3. [Online]. Available: https://www.researchgate.net/publication/313805219_Review_on_realization_of_AES_encryption_and_decryption_with_power_and_area_optimization
- [3] B. Zhang, G. Ma, X. Lu, and W. Xu, "Study on Hybrid Encryption Technology of Power Gateway Based on AES and RSA Algorithm," *2022 14th International Conference on Signal Processing Systems (ICSPS)*, Jiangsu, China, 2022, pp. 640-644. [Online]. Available: <https://www.computer.org/csdl/proceedings-article/icsp/2022/363100a640/IPTOEN5363m>
- [4] B. Harjito, H. M. Sukarno, and Winarno, "Performance Analysis of Kyber-DNA and RSA-Base64 Algorithms in Symmetric Key-Exchange Protocol," *2024 Ninth International Conference on Informatics and Computing (ICIC)*, Medan, Indonesia, 2024, pp. 1-6. [Online]. Available: <https://icic-aptikom.org/2024/wp-content/uploads/2024/10/Program-Book-ICIC-2024-Binder-2.pdf>



THANK YOU