Fraud Detection Using an Autoencoder and Variational Autoencoder

Group Members:

Name : Mayuresh Dindorkar
Roll No. : CS23MTECH14007

Name : Sanyam Kaul
Roll No. : CS23MTECH14011

Name : Shrenik Ganguli
Roll No. : CS23MTECH14014

Name : Morey Piyush Prabhakar
Roll No. : AI23MTECH14003

Indian Institute of Technology Hyderabad

Fraud Analytics (CS6890)

Index:

1. Problem Statement:

- Design a deep neural Autoencoder (AE) and Variational Autoencoder (VAE) to identify fraudulent transactions in the dataset of credit card transactions.

2. Dataset Description:

- The dataset contains 284315 ‘Normal’ (Genuine) and 492 ‘Fraud’ credit card transactions.

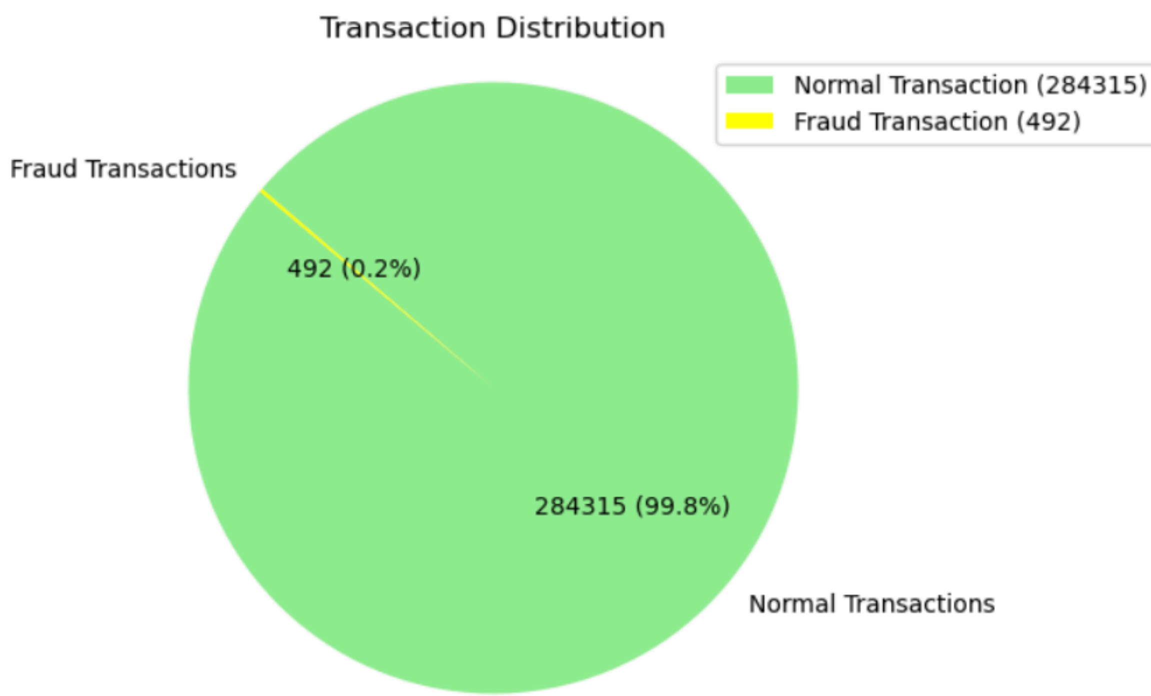


Figure 1: Pie chart depicting the class-wise distribution of transactions in the dataset

- The dataset comprises 30 features, with 28 being principal components, time difference and amount paid.

- The dataset doesn’t consist of any null values.

- We can observe that the dataset is highly skewed because the number of Normal credit card transactions is significantly larger compared to fraudulent transactions.

3. Auto-Encoder (AE):

- An autoencoder is an unsupervised deep neural network whose objective is to generate an output that closely resembles its input.

- It encodes the input into a lower dimensional space, usually called 'latent space', and then decodes it back to the original representation of the input.

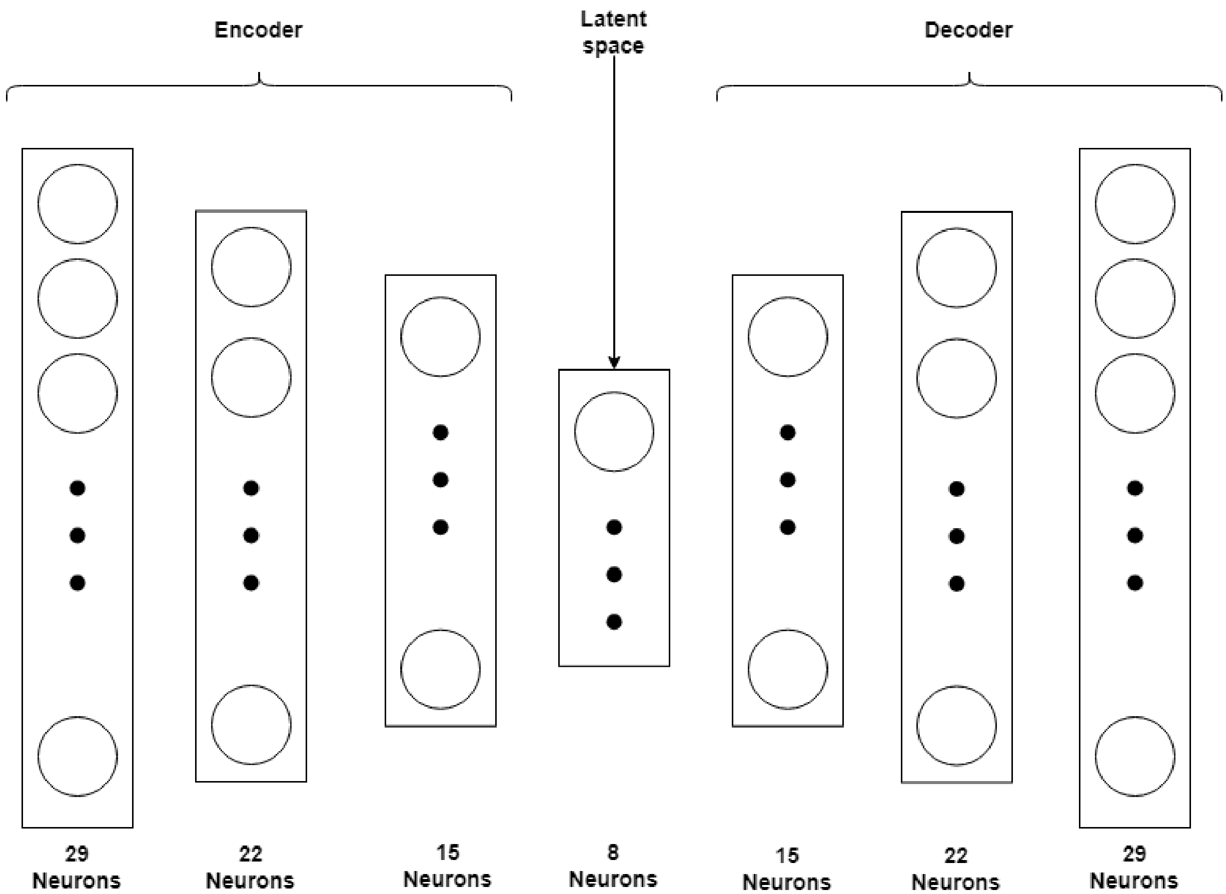


Figure 2: Architecture of Autoencoder used for classification of credit card transactions

- The number of neurons in the input layer of the autoencoder is equal to the number of features in the input dataset.

- Encoder: It encodes the input into a lower dimensional compressed latent representation.

- Decoder: It decodes the latent space representation into the original input space.

4. Variational Auto-Encoder (VAE):

- Variational autoencoder is a generative model that learns the probability distribution over the latent space.

- In both the training and inference stages, a Variational Autoencoder (VAE) leverages sampling from the latent space distribution it has learned.

- This sampling introduces randomness, enabling the model to produce varied outputs even when presented with the same input.

5. Results & Evaluation:

The dataset is segregated into Train, Validation and Test sets as below:

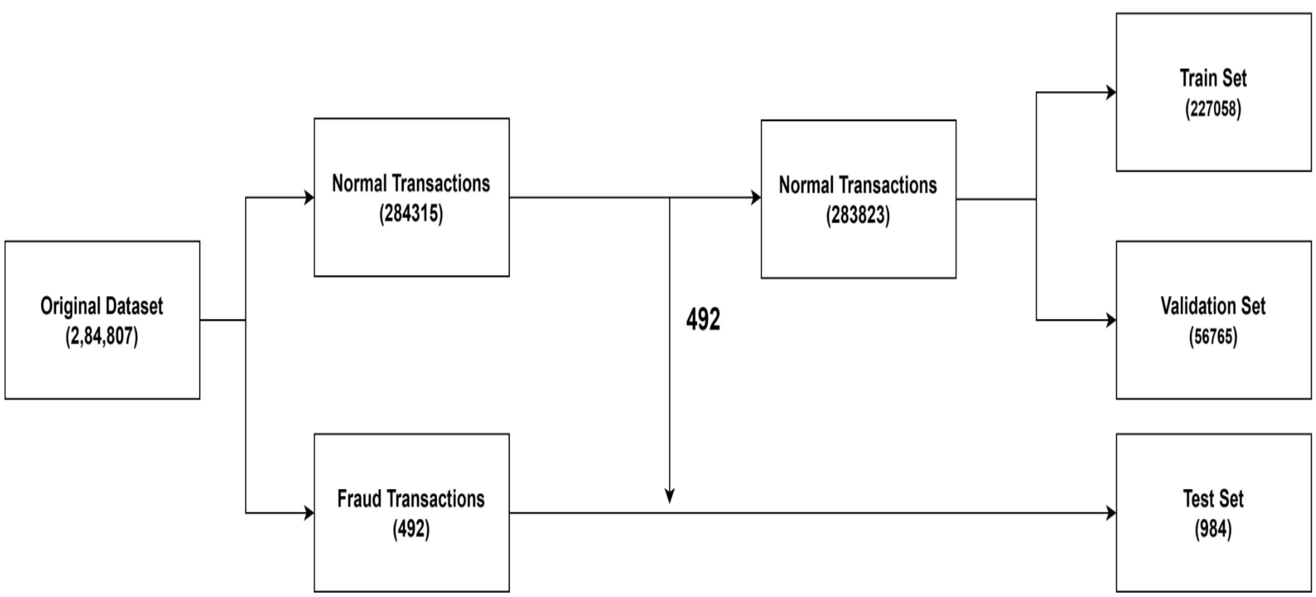


Figure 3: Dataset segregation into Train, Validation and Test set

5.1. Auto-Encoder results:

- Autoencoder is trained for ten epochs using Adam optimiser with 0.001 learning rate.
- The train loss graph is as below:

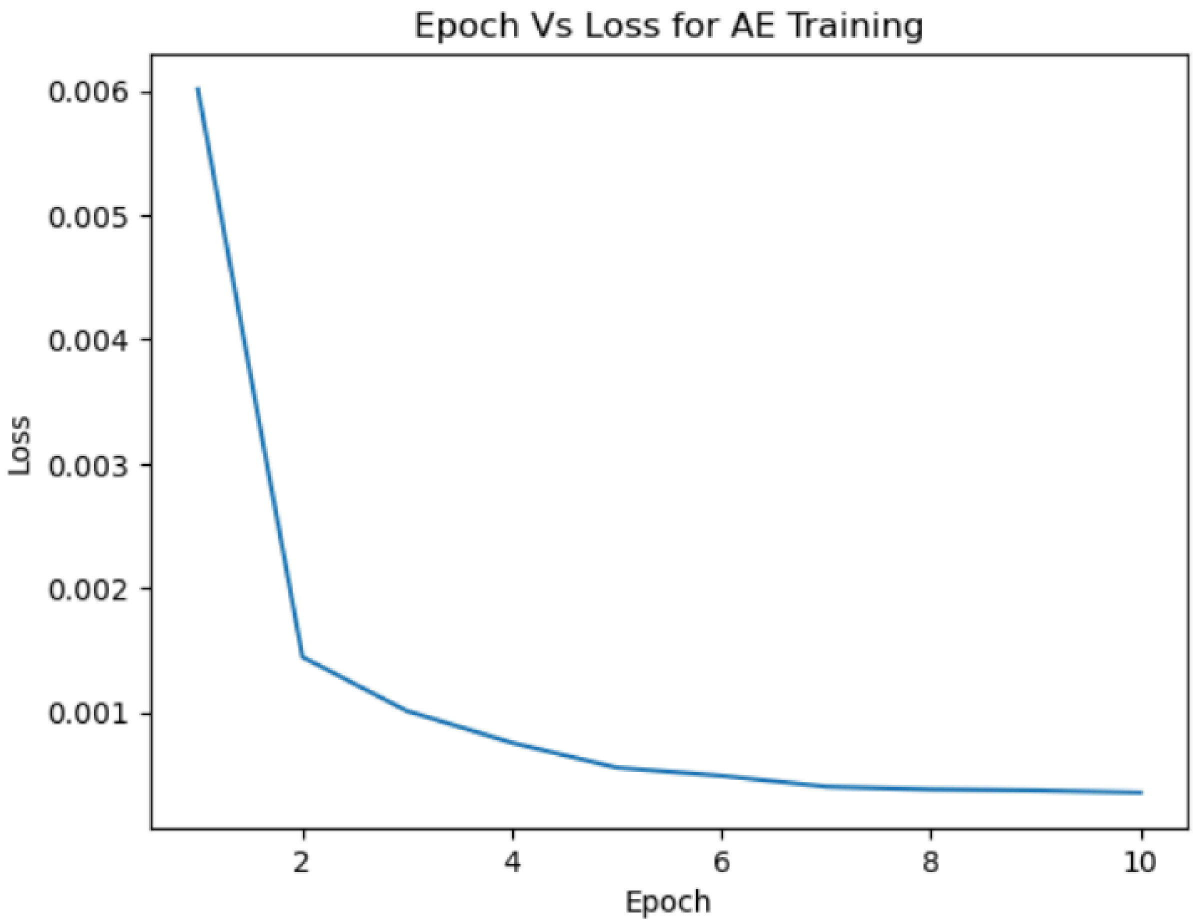


Figure 4: Training loss of Autoencoder

5.1.1. Mean Squared Error (MSE):

MSE on Validation Set	MSE on Test Set
0.00035	0.0030

5.1.2. Deciding threshold:

By observing the below graph, the threshold of 0.0036 is decided.

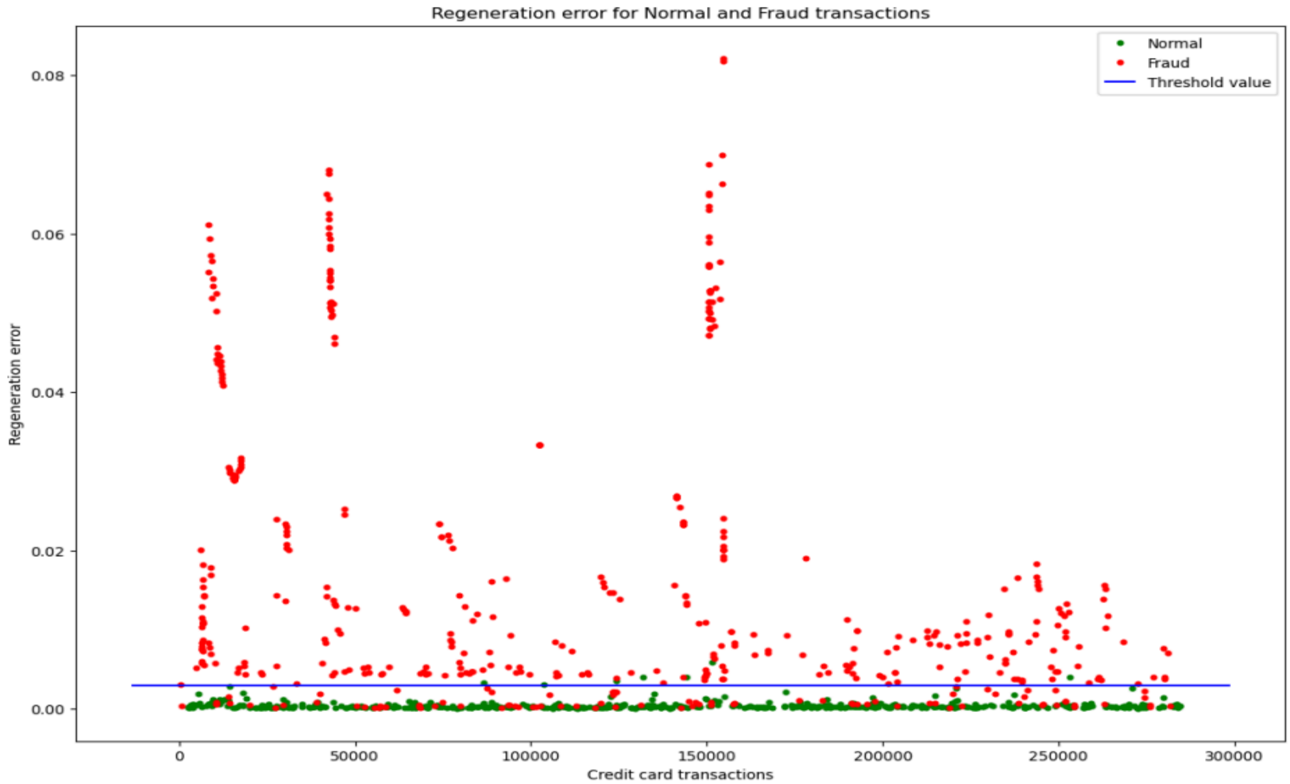
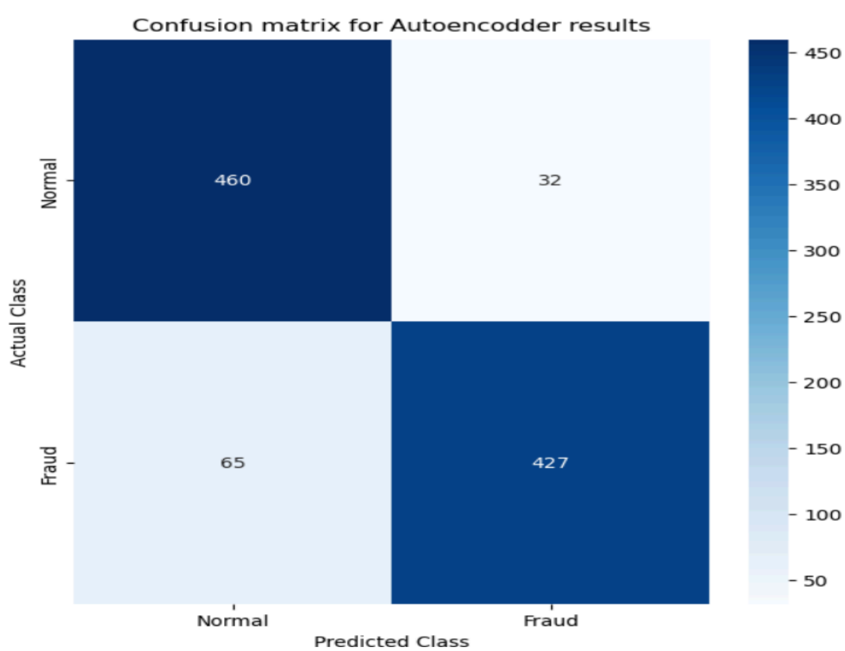


Figure 5: Transactions Vs Regeneration Error scatter plot

5.1.3. Confusion matrix:



5.1.4. ROC Curve:

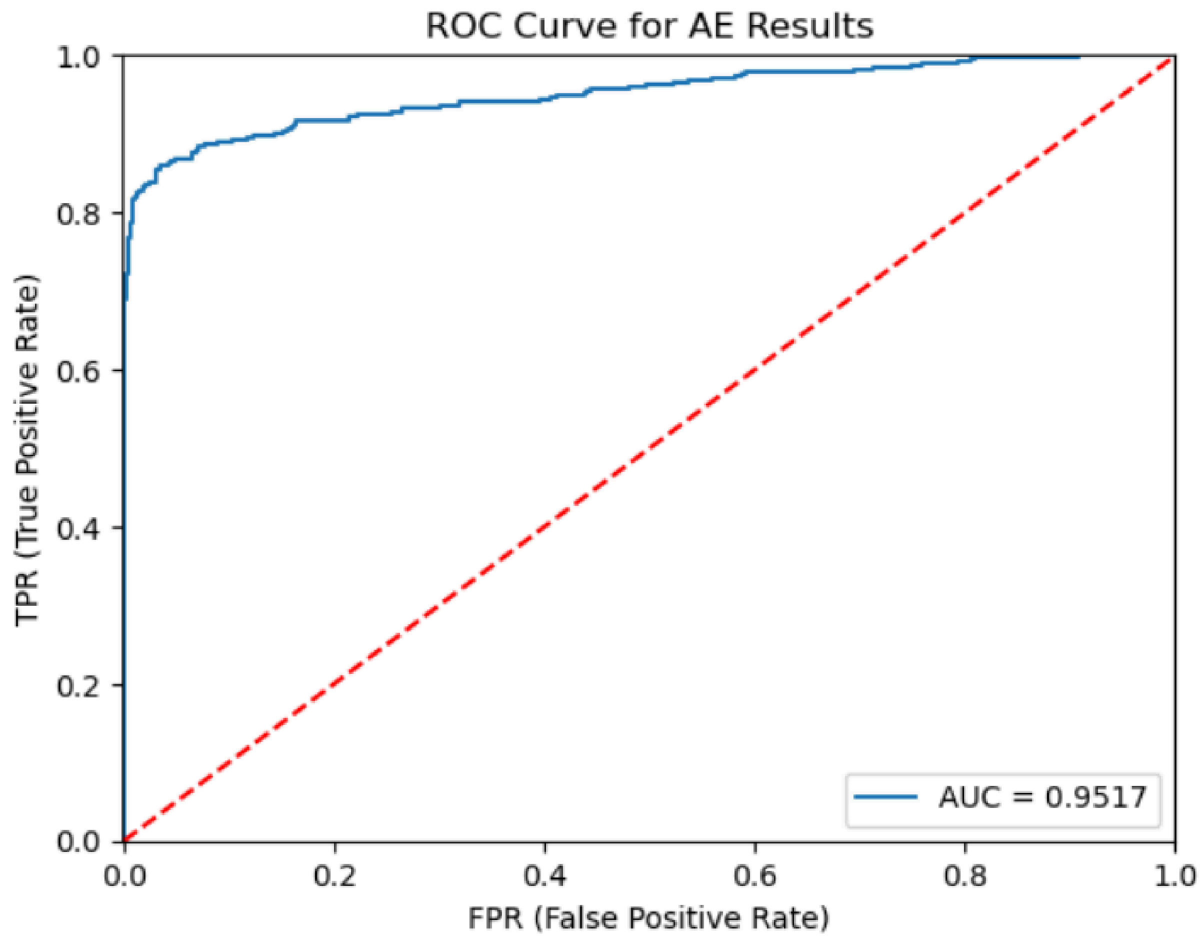


Figure 6: ROC curve for Autoencoder results on the Test set

5.1.5. Classification Report:

	precision	recall	f1-score	support
0	0.88	0.93	0.90	492
1	0.93	0.87	0.90	492
accuracy			0.90	984
macro avg	0.90	0.90	0.90	984
weighted avg	0.90	0.90	0.90	984

5.2. Variational Auto-Encoder (VAE):

- Variational autoencoder is also trained for ten epochs using Adam optimiser with 0.001 learning rate.
- The train loss graph is as below:

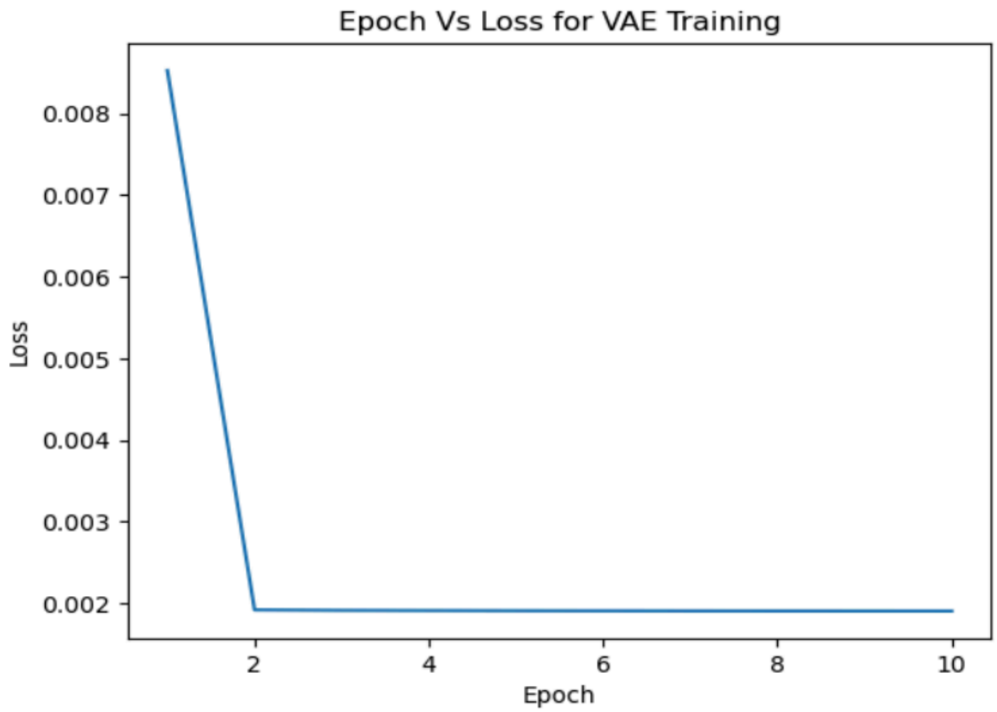


Figure 7: Training loss of Variational Autoencoder

5.2.1. Mean Squared Error (MSE):

MSE on Validation Set	MSE on Test Set
0.0019	0.0153

5.2.2. Deciding threshold:

By observing the below graph, the threshold of 0.0036 is decided.

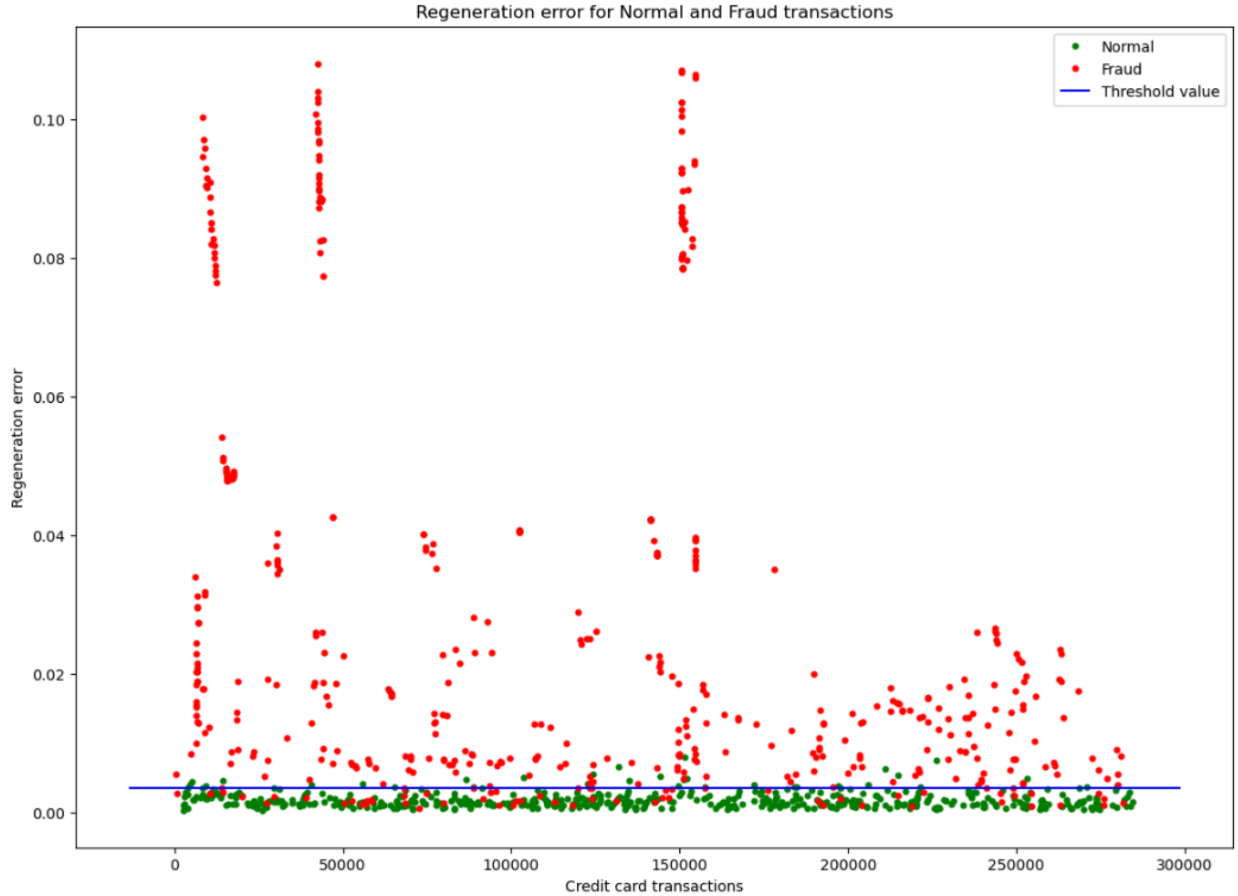
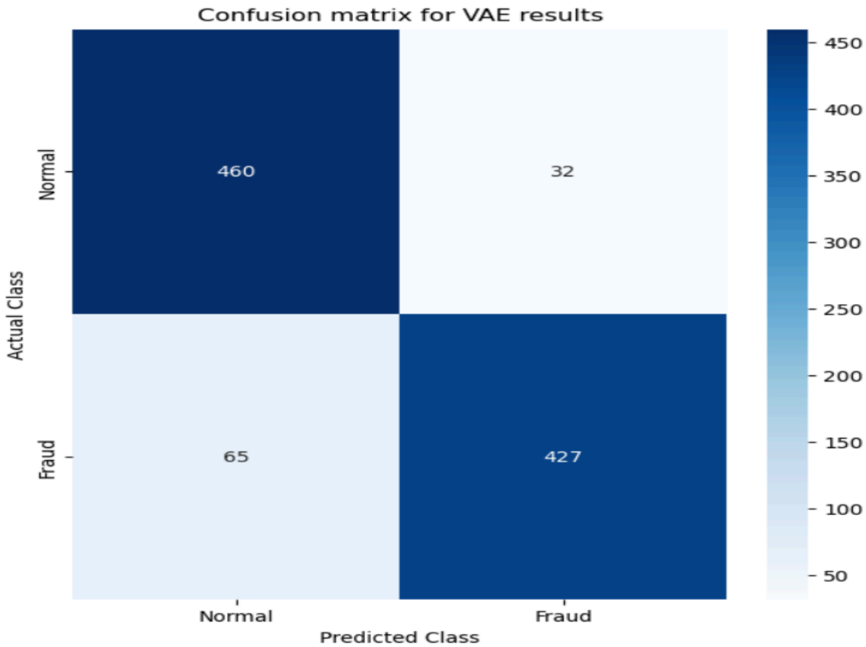


Figure 8: Transactions Vs Regeneration Error scatter plot

5.2.3. Confusion matrix:


Confusion matrix for VAE results
Actual Class
Normal
Fraud
460
32
65
427
Predicted Class
Normal
Fraud
450
400
350
300
250
200
150
100
50

5.2.4. ROC Curve:

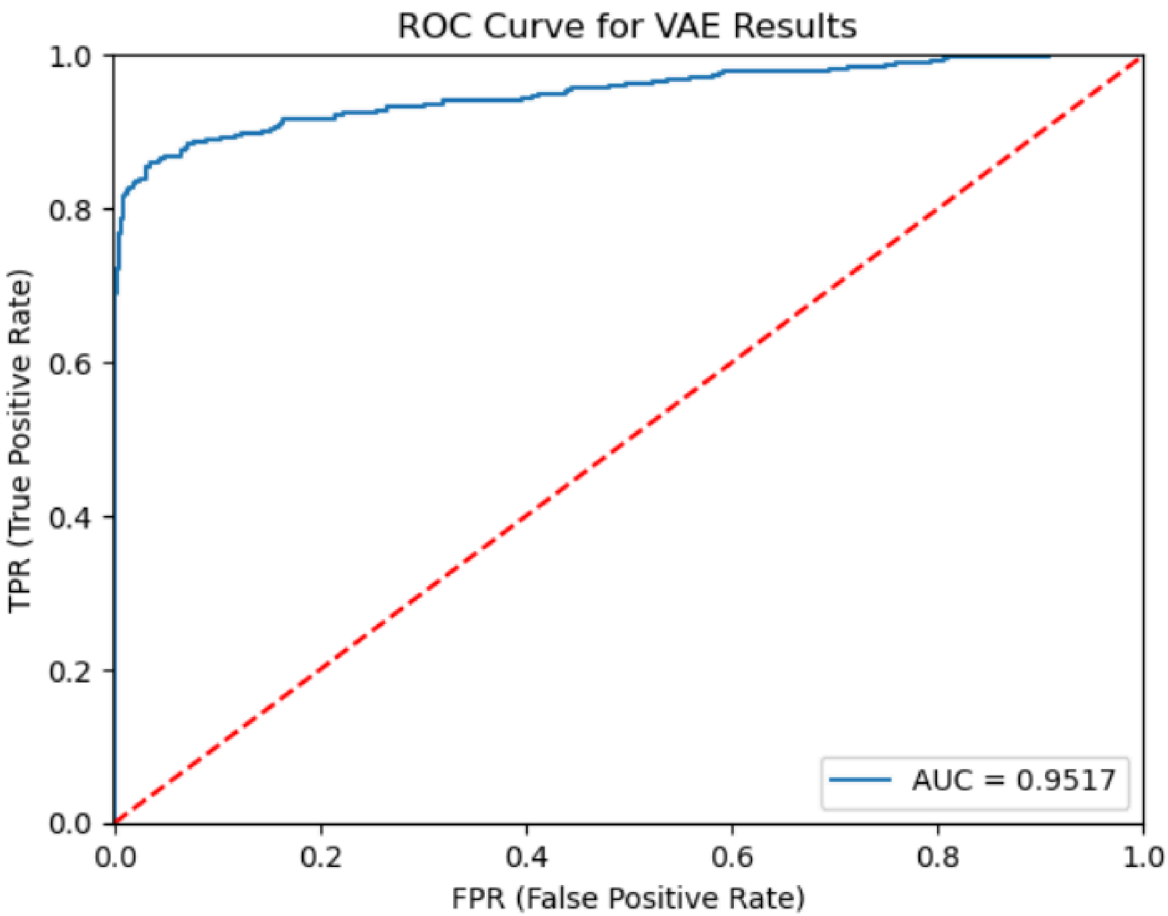


Figure 9: ROC curve for Variational Autoencoder results on the Test set

5.2.5. Classification Report:

	precision	recall	f1-score	support
0	0.87	0.95	0.91	492
1	0.94	0.86	0.90	492
accuracy			0.90	984
macro avg	0.91	0.90	0.90	984
weighted avg	0.91	0.90	0.90	984

6. Results:

- We can see from the observations that the optimal selection of threshold for VAE can outperform Autoencoder.

- For same value of the threshold, VAE has given a slightly better F1-Score for both classes compared to AE.

- Through VAE we have learned the probabilistic distribution of our data, which can be utilized for synthetic data generation.

Model	F1-Score for Class '0' on Test set	F1-Score for Class '1' on Test set
AE	0.90	0.90
VAE	0.91	0.90