

```
In [1]: !pip install tensorflow
```

Requirement already satisfied: tensorflow in c:\programdata\anaconda3\lib\site-packages (2.17.0)
Requirement already satisfied: tensorflow-intel==2.17.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow) (2.17.0)
Requirement already satisfied: absl-py>=1.0.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (2.1.0)
Requirement already satisfied: astunparse>=1.6.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (1.6.3)
Requirement already satisfied: flatbuffers>=24.3.25 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (24.3.25)
Requirement already satisfied: gast!=0.5.0,!0.5.1,!0.5.2,>=0.2.1 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (0.6.0)
Requirement already satisfied: google-pasta>=0.1.1 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (0.2.0)
Requirement already satisfied: h5py>=3.10.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (3.11.0)
Requirement already satisfied: libclang>=13.0.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (18.1.1)
Requirement already satisfied: ml-dtypes<0.5.0,>=0.3.1 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (0.4.0)
Requirement already satisfied: opt-einsum>=2.3.2 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (3.3.0)
Requirement already satisfied: packaging in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (23.1)
Requirement already satisfied: protobuf!=4.21.0,!4.21.1,!4.21.2,!4.21.3,!4.21.4,!4.21.5,<5.0.0dev,>=3.20.3 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (3.20.3)
Requirement already satisfied: requests<3,>=2.21.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (2.31.0)
Requirement already satisfied: setuptools in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (68.2.2)
Requirement already satisfied: six>=1.12.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (1.16.0)
Requirement already satisfied: termcolor>=1.1.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (2.4.0)
Requirement already satisfied: typing-extensions>=3.6.6 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (4.9.0)
Requirement already satisfied: wrapt>=1.11.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (1.14.1)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (1.66.0)
Requirement already satisfied: tensorboard<2.18,>=2.17 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (2.17.1)

Requirement already satisfied: keras>=3.2.0 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (3.5.0)

Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (0.31.0)

Requirement already satisfied: numpy<2.0.0,>=1.23.5 in c:\programdata\anaconda3\lib\site-packages (from tensorflow-intel==2.17.0->tensorflow) (1.26.4)

Requirement already satisfied: wheel<1.0,>=0.23.0 in c:\programdata\anaconda3\lib\site-packages (from astunparse>=1.6.0->tensorflow-intel==2.17.0->tensorflow) (0.41.2)

Requirement already satisfied: rich in c:\programdata\anaconda3\lib\site-packages (from keras>=3.2.0->tensorflow-intel==2.17.0->tensorflow) (13.3.5)

Requirement already satisfied: namex in c:\programdata\anaconda3\lib\site-packages (from keras>=3.2.0->tensorflow-intel==2.17.0->tensorflow) (0.0.8)

Requirement already satisfied: optree in c:\programdata\anaconda3\lib\site-packages (from keras>=3.2.0->tensorflow-intel==2.17.0->tensorflow) (0.12.1)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\programdata\anaconda3\lib\site-packages (from requests<3,>=2.21.0->tensorflow-intel==2.17.0->tensorflow) (2.0.4)

Requirement already satisfied: idna<4,>=2.5 in c:\programdata\anaconda3\lib\site-packages (from requests<3,>=2.21.0->tensorflow-intel==2.17.0->tensorflow) (3.4)

Requirement already satisfied: urllib3<3,>=1.21.1 in c:\programdata\anaconda3\lib\site-packages (from requests<3,>=2.21.0->tensorflow-intel==2.17.0->tensorflow) (2.0.7)

Requirement already satisfied: certifi>=2017.4.17 in c:\programdata\anaconda3\lib\site-packages (from requests<3,>=2.21.0->tensorflow-intel==2.17.0->tensorflow) (2024.2.2)

Requirement already satisfied: markdown>=2.6.8 in c:\programdata\anaconda3\lib\site-packages (from tensorboard<2.18,>=2.17->tensorflow-intel==2.17.0->tensorflow) (3.4.1)

Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in c:\programdata\anaconda3\lib\site-packages (from tensorboard<2.18,>=2.17->tensorflow-intel==2.17.0->tensorflow) (0.7.2)

Requirement already satisfied: werkzeug>=1.0.1 in c:\programdata\anaconda3\lib\site-packages (from tensorboard<2.18,>=2.17->tensorflow-intel==2.17.0->tensorflow) (2.2.3)

Requirement already satisfied: MarkupSafe>=2.1.1 in c:\programdata\anaconda3\lib\site-packages (from werkzeug>=1.0.1->tensorboard<2.18,>=2.17->tensorflow-intel==2.17.0->tensorflow) (2.1.3)

Requirement already satisfied: markdown-it-py<3.0.0,>=2.2.0 in c:\programdata\anaconda3\lib\site-packages (from rich->keras>=3.2.0->tensorflow-intel==2.17.0->tensorflow) (2.2.0)

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in c:\programdata\anaconda3\lib\site-packages (from rich->keras>=3.2.0->tensorflow-intel==2.17.0->tensorflow) (2.15.1)

Requirement already satisfied: mdurl~0.1 in c:\programdata\anaconda3\lib\site-packages (from markdown-it-py<3.0.0,>=2.2.0->rich->keras>=3.2.0->tensorflow-intel==2.17.0->tensorflow) (0.1.0)

DEPRECATION: Loading egg at c:\programdata\anaconda3\lib\site-packages\vboxapi-1.0-py3.11.egg is deprecated. pip 24.3 will enforce this behaviour change. A possible replacement is to use pip for package installation.. Discussion can be found at <https://github.com/pypa/pip/issues/12330>

```
In [2]: import pandas as pd
import numpy as np
import tensorflow as tf
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler
from sklearn.metrics import confusion_matrix, recall_score, accuracy_score, precision_score

RANDOM_SEED = 2021
TEST_PCT = 0.3
LABELS = ["Normal", "Fraud"]
```

```
In [3]: dataset = pd.read_csv("creditcard.csv")
```

```
In [4]: #check for any null values
print("Any nulls in the dataset", dataset.isnull().values.any())
print('-----')
print("No. of unique labels", len(dataset['Class'].unique()))
print("Label values", dataset.Class.unique())

#0 is for normal credit card transaction
#1 is for fraudulent credit card transaction
print('-----')
print("Break down of Normal and Fraud Transactions")
print(pd.value_counts(dataset['Class'], sort=True))
```

Any nulls in the dataset False

No. of unique labels 2

Label values [0 1]

Break down of Normal and Fraud Transactions

Class

0 284315

1 492

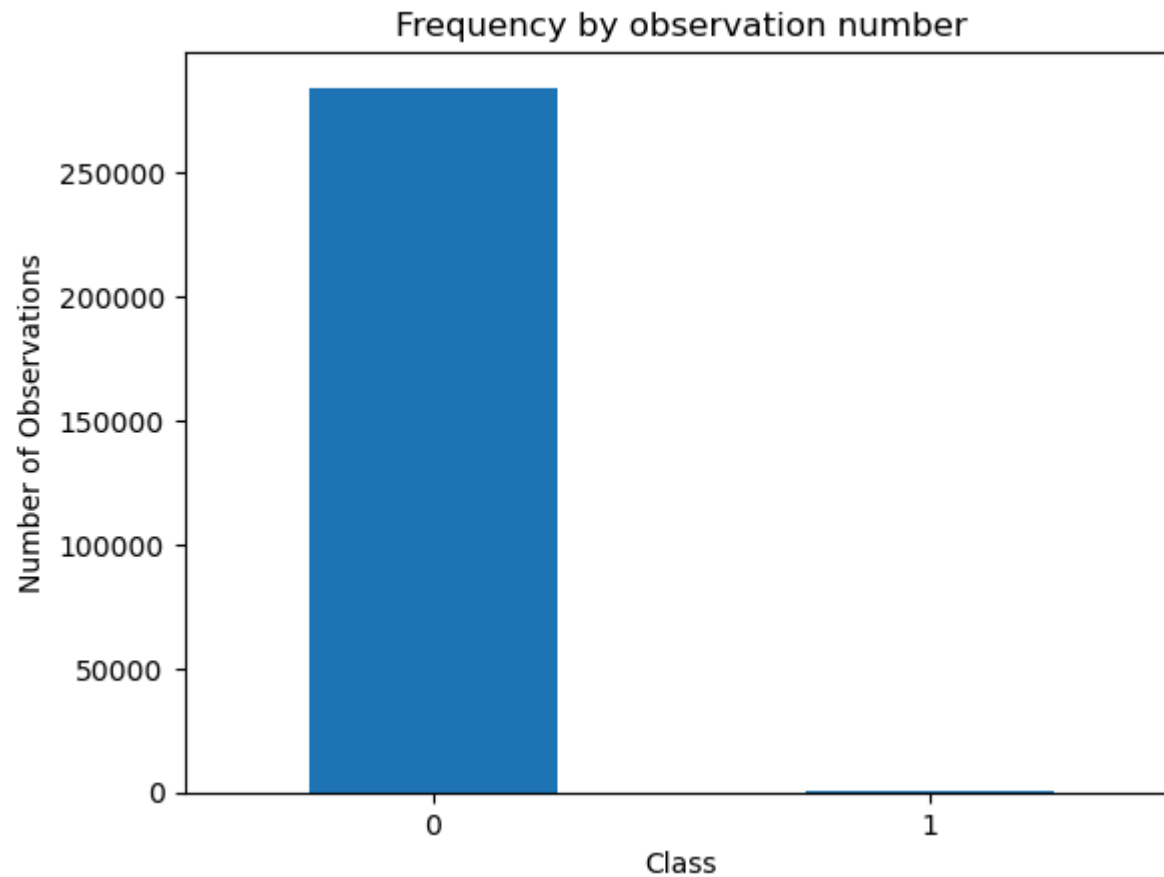
Name: count, dtype: int64

```
C:\Users\Devang Tilgule\AppData\Local\Temp\ipykernel_12560\3382237470.py:11: FutureWarning: pandas.value_counts is deprecated and will be removed in a future version. Use pd.Series(obj).value_counts() instead.  
print(pd.value_counts(dataset['Class'],sort=True))
```

```
In [5]: #visualizing the imbalanced dataset  
count_classes = pd.value_counts(dataset['Class'],sort=True)  
count_classes.plot(kind='bar',rot=0)  
plt.xticks(range(len(dataset['Class'].unique())),dataset.Class.unique())  
plt.title("Frequency by observation number")  
plt.xlabel("Class")  
plt.ylabel("Number of Observations")
```

```
C:\Users\Devang Tilgule\AppData\Local\Temp\ipykernel_12560\3824744727.py:2: FutureWarning: pandas.value_counts is deprecated and will be removed in a future version. Use pd.Series(obj).value_counts() instead.  
count_classes = pd.value_counts(dataset['Class'],sort=True)
```

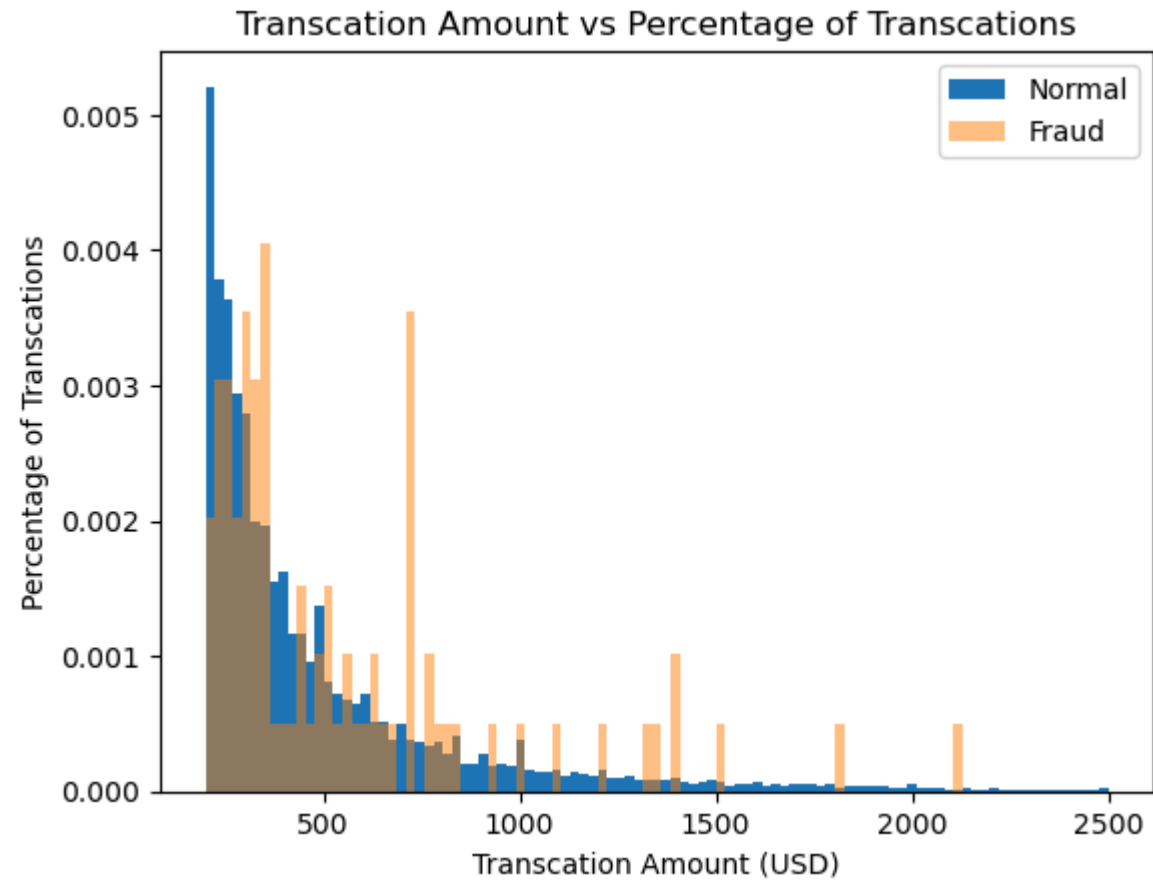
```
Out[5]: Text(0, 0.5, 'Number of Observations')
```



```
In [6]: #Save the normal and fradulent transctions in seperate dataframe
normal_dataset = dataset[dataset.Class == 0]
fraud_dataset = dataset[dataset.Class == 1]

#Visualize transcation amounts for normal and fraudulent transctions
bins = np.linspace(200,2500,100)
plt.hist(normal_dataset.Amount,bins=bins,alpha=1,density=True,label='Normal')
plt.hist(fraud_dataset.Amount,bins=bins,alpha=0.5,density=True,label='Fraud')
plt.legend(loc='upper right')
plt.title("Transcation Amount vs Percentage of Transcations")
plt.xlabel("Transcation Amount (USD)")
```

```
plt.ylabel("Percentage of Transactions")  
plt.show()
```



```
In [7]: dataset
```

Out[7]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	...	-0.018307	0.2778
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	...	-0.225775	-0.6386
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	...	0.247998	0.7716
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	...	-0.108300	0.0052
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	...	-0.009431	0.7982
...
284802	172786.0	-11.881118	10.071785	-9.834783	-2.066656	-5.364473	-2.606837	-4.918215	7.305334	1.914428	...	0.213454	0.1118
284803	172787.0	-0.732789	-0.055080	2.035030	-0.738589	0.868229	1.058415	0.024330	0.294869	0.584800	...	0.214205	0.9243
284804	172788.0	1.919565	-0.301254	-3.249640	-0.557828	2.630515	3.031260	-0.296827	0.708417	0.432454	...	0.232045	0.5782
284805	172788.0	-0.240440	0.530483	0.702510	0.689799	-0.377961	0.623708	-0.686180	0.679145	0.392087	...	0.265245	0.8000
284806	172792.0	-0.533413	-0.189733	0.703337	-0.506271	-0.012546	-0.649617	1.577006	-0.414650	0.486180	...	0.261057	0.6430

284807 rows × 31 columns



```
In [8]: sc = StandardScaler()
dataset['Time'] = sc.fit_transform(dataset['Time'].values.reshape(-1,1))
dataset['Amount'] = sc.fit_transform(dataset['Amount'].values.reshape(-1,1))
```

```
In [9]: raw_data = dataset.values
#The last element contains if the transaction is normal which is represented by 0 and if fraud then 1
labels = raw_data[:, -1]

#The other data points are the electrocardiogram data
data = raw_data[:, 0:-1]

train_data, test_data, train_labels, test_labels = train_test_split(data, labels, test_size = 0.2, random_state = 2021)
```



```
In [10]: min_val = tf.reduce_min(train_data)
max_val = tf.reduce_max(train_data)

train_data = (train_data - min_val) / (max_val - min_val)
test_data = (test_data - min_val) / (max_val - min_val)

train_data = tf.cast(train_data,tf.float32)
test_data = tf.cast(test_data,tf.float32)
```

```
In [11]: train_labels = train_labels.astype(bool)
test_labels = test_labels.astype(bool)

#Creating normal and fraud datasets
normal_train_data = train_data[~train_labels]
normal_test_data = test_data[~test_labels]

fraud_train_data = train_data[train_labels]
fraud_test_data = test_data[test_labels]
print("No. of records in Fraud Train Data=",len(fraud_train_data))
print("No. of records in Normal Train Data=",len(normal_train_data))
print("No. of records in Fraud Test Data=",len(fraud_test_data))
print("No. of records in Normal Test Data=",len(normal_test_data))
```

```
No. of records in Fraud Train Data= 389
No. of records in Normal Train Data= 227456
No. of records in Fraud Test Data= 103
No. of records in Normal Test Data= 56859
```

```
In [12]: nb_epoch = 50
batch_size = 64
input_dim = normal_train_data.shape[1]
#num of columns,30
encoding_dim = 14
hidden_dim1 = int(encoding_dim / 2)
hidden_dim2 = 4
learning_rate = 1e-7
```

```
In [13]: #input layer
input_layer = tf.keras.layers.Input(shape=(input_dim,))
```

```

#Encoder
encoder = tf.keras.layers.Dense(encoding_dim,activation="tanh",activity_regularizer = tf.keras.regularizers.l2(learning_rate))
encoder = tf.keras.layers.Dropout(0.2)(encoder)
encoder = tf.keras.layers.Dense(hidden_dim1,activation='relu')(encoder)
encoder = tf.keras.layers.Dense(hidden_dim2,activation=tf.nn.leaky_relu)(encoder)

#Decoder
decoder = tf.keras.layers.Dense(hidden_dim1,activation='relu')(encoder)
decoder = tf.keras.layers.Dropout(0.2)(decoder)
decoder = tf.keras.layers.Dense(encoding_dim,activation='relu')(decoder)
decoder = tf.keras.layers.Dense(input_dim,activation='tanh')(decoder)

#Autoencoder
autoencoder = tf.keras.Model(inputs = input_layer,outputs = decoder)
autoencoder.summary()

```

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 30)	0
dense (Dense)	(None, 14)	434
dropout (Dropout)	(None, 14)	0
dense_1 (Dense)	(None, 7)	105
dense_2 (Dense)	(None, 4)	32
dense_3 (Dense)	(None, 7)	35
dropout_1 (Dropout)	(None, 7)	0
dense_4 (Dense)	(None, 14)	112
dense_5 (Dense)	(None, 30)	450

Total params: 1,168 (4.56 KB)

Trainable params: 1,168 (4.56 KB)

Non-trainable params: 0 (0.00 B)

```
In [15]: cp = tf.keras.callbacks.ModelCheckpoint(filepath="autoencoder_fraud.keras",mode='min',monitor='val_loss',verbose=2,save_best_c
#Define our early stopping
early_stop = tf.keras.callbacks.EarlyStopping(
    monitor='val_loss',
    min_delta=0.0001,
    patience=10,
    verbose=11,
    mode='min',
    restore_best_weights=True
)
```

```
In [16]: autoencoder.compile(metrics=['accuracy'],loss= 'mean_squared_error',optimizer='adam')
```

```
In [17]: history = autoencoder.fit(normal_train_data,normal_train_data,epochs = nb_epoch,
    batch_size = batch_size,shuffle = True,
    validation_data = (test_data,test_data),
    verbose=1,
    callbacks = [cp,early_stop]).history
```

Epoch 1/50
3539/3554 ————— 0s 2ms/step - accuracy: 0.0296 - loss: 0.0177
Epoch 1: val_loss improved from inf to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 12s 3ms/step - accuracy: 0.0296 - loss: 0.0177 - val_accuracy: 0.0190 - val_loss: 2.2905e-05
Epoch 2/50
3554/3554 ————— 0s 2ms/step - accuracy: 0.0602 - loss: 1.9000e-05
Epoch 2: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 3ms/step - accuracy: 0.0602 - loss: 1.9000e-05 - val_accuracy: 0.0190 - val_loss: 2.2504e-05
Epoch 3/50
3541/3554 ————— 0s 2ms/step - accuracy: 0.0616 - loss: 1.9284e-05
Epoch 3: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 10s 3ms/step - accuracy: 0.0616 - loss: 1.9285e-05 - val_accuracy: 0.0254 - val_loss: 2.1493e-05
Epoch 4/50
3538/3554 ————— 0s 2ms/step - accuracy: 0.0820 - loss: 1.9194e-05
Epoch 4: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 2ms/step - accuracy: 0.0821 - loss: 1.9193e-05 - val_accuracy: 0.0323 - val_loss: 1.9062e-05
Epoch 5/50
3545/3554 ————— 0s 2ms/step - accuracy: 0.1242 - loss: 1.8436e-05
Epoch 5: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 2ms/step - accuracy: 0.1242 - loss: 1.8435e-05 - val_accuracy: 0.0485 - val_loss: 1.8428e-05
Epoch 6/50
3549/3554 ————— 0s 2ms/step - accuracy: 0.1392 - loss: 1.8057e-05
Epoch 6: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 3ms/step - accuracy: 0.1392 - loss: 1.8057e-05 - val_accuracy: 0.2744 - val_loss: 1.7884e-05
Epoch 7/50
3552/3554 ————— 0s 2ms/step - accuracy: 0.1980 - loss: 1.7659e-05
Epoch 7: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 2ms/step - accuracy: 0.1980 - loss: 1.7659e-05 - val_accuracy: 0.2855 - val_loss: 1.7141e-05
Epoch 8/50
3527/3554 ————— 0s 2ms/step - accuracy: 0.2607 - loss: 1.7131e-05
Epoch 8: val_loss did not improve from 0.00002
3554/3554 ————— 9s 3ms/step - accuracy: 0.2607 - loss: 1.7130e-05 - val_accuracy: 0.2493 - val_loss: 1.7996e-05
Epoch 9/50
3546/3554 ————— 0s 2ms/step - accuracy: 0.2699 - loss: 1.6883e-05
Epoch 9: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 3ms/step - accuracy: 0.2699 - loss: 1.6883e-05 - val_accuracy: 0.3591 - val_loss: 1.6375e-05
Epoch 10/50
3531/3554 ————— 0s 2ms/step - accuracy: 0.2938 - loss: 1.6583e-05
Epoch 10: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras
3554/3554 ————— 9s 3ms/step - accuracy: 0.2938 - loss: 1.6582e-05 - val_accuracy: 0.3642 - val_loss: 1.6047e-05
Epoch 11/50

3532/3554 — 0s 2ms/step - accuracy: 0.3047 - loss: 1.6730e-05

Epoch 11: val_loss improved from 0.00002 to 0.00002, saving model to autoencoder_fraud.keras

3554/3554 — 9s 3ms/step - accuracy: 0.3047 - loss: 1.6727e-05 - val_accuracy: 0.3308 - val_loss: 1.6004e-05

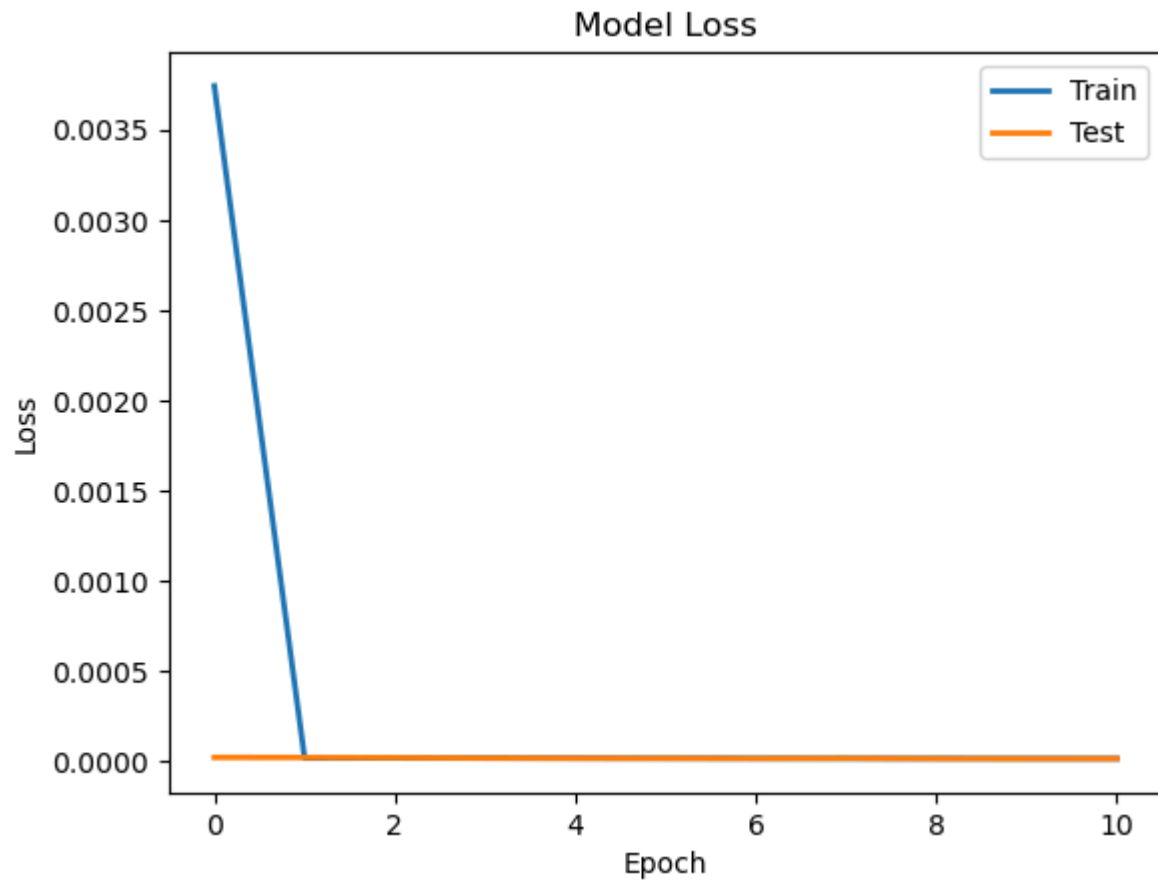
Epoch 11: early stopping

Restoring model weights from the end of the best epoch: 1.

```
In [18]: plt.plot(history['loss'],linewidth = 2,label = 'Train')
plt.plot(history['val_loss'],linewidth = 2,label = 'Test')
plt.legend(loc='upper right')
plt.title('Model Loss')
plt.ylabel('Loss')
plt.xlabel('Epoch')

#plt.ylim(ymin=0.70,ymax=1)

plt.show()
```



```
In [19]: test_x_predictions = autoencoder.predict(test_data)
mse = np.mean(np.power(test_data - test_x_predictions, 2),axis = 1)
error_df = pd.DataFrame({'Reconstruction_error':mse,
                        'True_class':test_labels})
```

1781/1781 ————— 3s 2ms/step

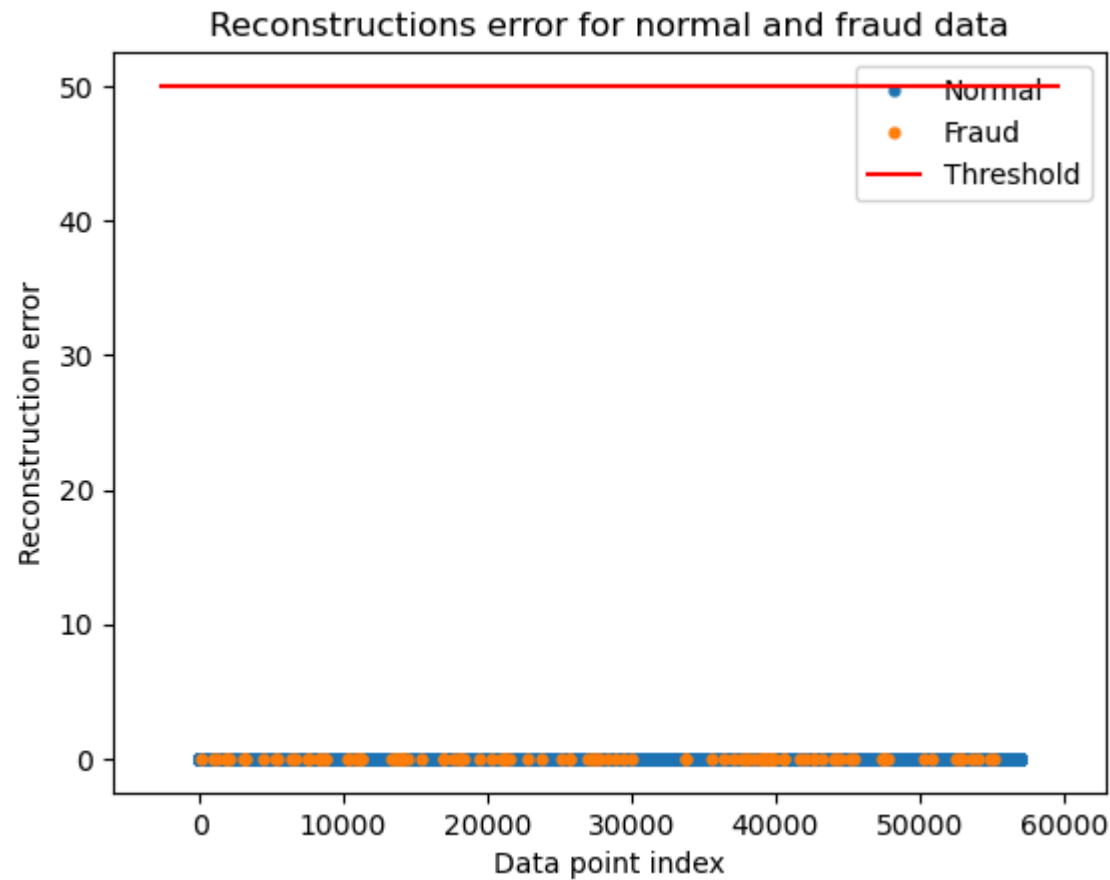
```
In [20]: threshold_fixed = 50
groups = error_df.groupby('True_class')
fig,ax = plt.subplots()

for name,group in groups:
    ax.plot(group.index,group.Reconstruction_error,marker='o',ms = 3.5,linestyle='',
```

```

        label = "Fraud" if name==1 else "Normal")
ax.hlines(threshold_fixed,ax.get_xlim()[0],ax.get_xlim()[1],colors="r",zorder=100,label="Threshold")
ax.legend()
plt.title("Reconstructions error for normal and fraud data")
plt.ylabel("Reconstruction error")
plt.xlabel("Data point index")
plt.show()

```



```

In [21]: threshold_fixed = 52
pred_y = [1 if e > threshold_fixed else 0
          for e in
            error_df.Reconstruction_error.values]
error_df['pred'] = pred_y

```

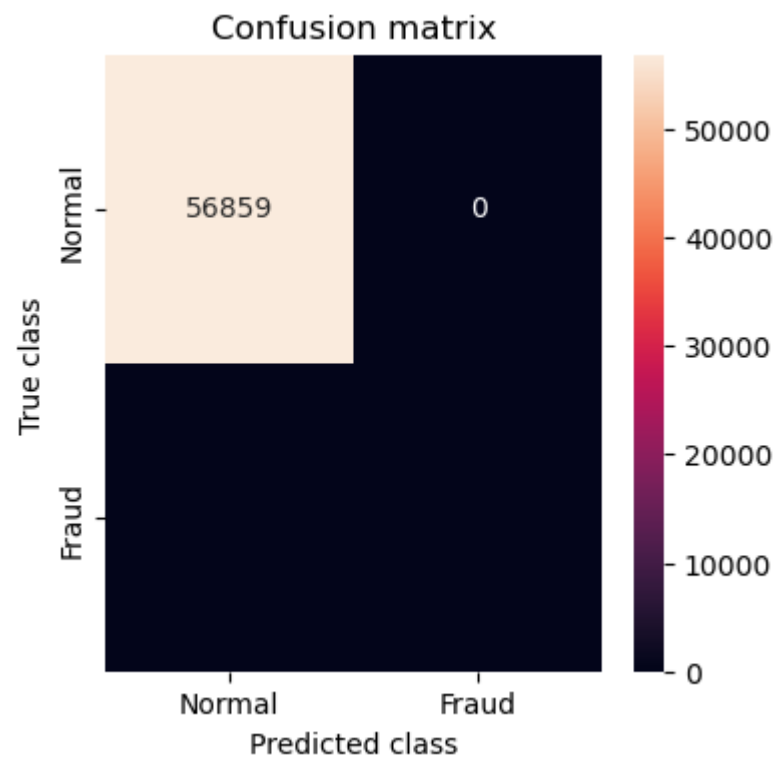
```

conf_matrix = confusion_matrix(error_df.True_class,pred_y)

plt.figure(figsize = (4,4))
sns.heatmap(conf_matrix,xticklabels = LABELS,yticklabels = LABELS,annot = True,fmt="d")
plt.title("Confusion matrix")
plt.ylabel("True class")
plt.xlabel("Predicted class")
plt.show()

#Print Accuracy,Precision and Recall
print("Accuracy :",accuracy_score(error_df['True_class'],error_df['pred']))
print("Recall :",recall_score(error_df['True_class'],error_df['pred']))
print("Precision :",precision_score(error_df['True_class'],error_df['pred']))

```



Accuracy : 0.9981917769741231
 Recall : 0.0
 Precision : 0.0


```
C:\ProgramData\anaconda3\Lib\site-packages\sklearn\metrics\_classification.py:1344: UndefinedMetricWarning: Precision is ill-defined and being set to 0.0 due to no predicted samples. Use `zero_division` parameter to control this behavior.  
  _warn_prf(average, modifier, msg_start, len(result))
```

In []: