Assignment 4

1. Name: Vaibhav Bichave
2. Batch: P-10
3. Roll No.: 43209

Problem Statement :

ECG Anomaly detection using Autoencoders

a. Import required libraries

In [10]:

```
import numpy as np
import pandas as pd
import tensorflow as tf
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score
from tensorflow.keras.optimizers import Adam
from sklearn.preprocessing import MinMaxScaler
from tensorflow.keras import Model, Sequential
from tensorflow.keras.layers import Dense, Dropout
from sklearn.model_selection import train_test_split
from tensorflow.keras.losses import MeanSquaredLogarithmicError
```

b. Upload / access the dataset

In [11]:

```
PATH_TO_DATA = 'http://storage.googleapis.com/download.tensorflow.org/data/ecg.csv'
data = pd.read_csv(PATH_TO_DATA, header=None)
data.head()
```

Out[11]:

	0	1	2	3	4	5	6	7	
0	-0.112522	-2.827204	-3.773897	-4.349751	-4.376041	-3.474986	-2.181408	-1.818286	-1.250
1	-1.100878	-3.996840	-4.285843	-4.506579	-4.022377	-3.234368	-1.566126	-0.992258	-0.754
2	-0.567088	-2.593450	-3.874230	-4.584095	-4.187449	-3.151462	-1.742940	-1.490659	-1.183
3	0.490473	-1.914407	-3.616364	-4.318823	-4.268016	-3.881110	-2.993280	-1.671131	-1.333
4	0.800232	-0.874252	-2.384761	-3.973292	-4.338224	-3.802422	-2.534510	-1.783423	-1.594

5 rows × 141 columns

Finding shape of the dataset

In [2]:

```
data.shape
```

Out[2]:

```
(4998, 141)
```

Splitting training and testing dataset

In [3]:

```
features = data.drop(140, axis=1)
target = data[140]
x_train, x_test, y_train, y_test = train_test_split(
    features, target, test_size=0.2, stratify=target
)
train_index = y_train[y_train == 1].index
train_data = x_train.loc[train_index]
```

Scaling the data using MinMaxScaler

In [12]:

```
min_max_scaler = MinMaxScaler(feature_range=(0, 1))
x_train_scaled = min_max_scaler.fit_transform(train_data.copy())
x_test_scaled = min_max_scaler.transform(x_test.copy())
```

c. Encoder converts it into latent representation

d. Decoder networks convert it back to the original input

Creating autoencoder subclass by extending Model class from keras

In [13]:

```
class AutoEncoder(Model):
  def __init__(self, output_units, ldim=8):
    super().__init__()
    self.encoder = Sequential([
      Dense(64, activation='relu'),
      Dropout(0.1),
      Dense(32, activation='relu'),
      Dropout(0.1),
      Dense(16, activation='relu'),
      Dropout(0.1),
      Dense(ldim, activation='relu')
    ])
    self.decoder = Sequential([
      Dense(16, activation='relu'),
      Dropout(0.1),
      Dense(32, activation='relu'),
      Dropout(0.1),
      Dense(64, activation='relu'),
      Dropout(0.1),
      Dense(output_units, activation='sigmoid')
    ])

  def call(self, inputs):
    encoded = self.encoder(inputs)
    decoded = self.decoder(encoded)
    return decoded
```

e. Compile the models with Optimizer, Loss, and Evaluation Metrics

Model configuration

In [14]:

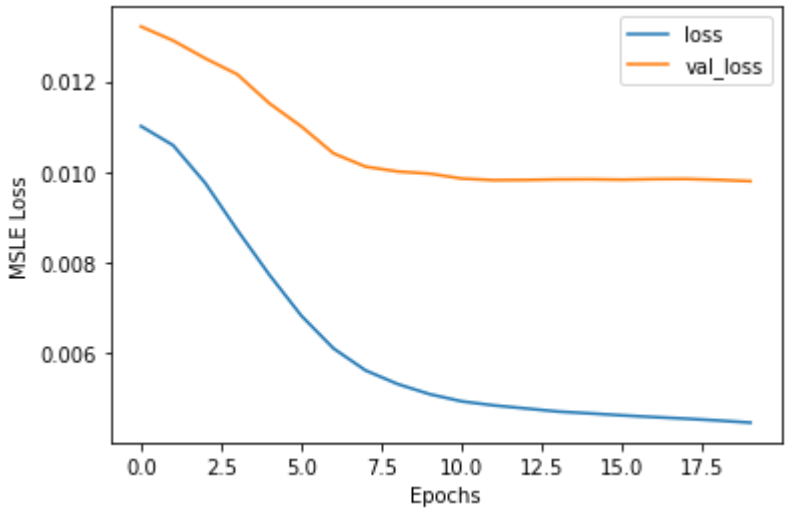
```
model = AutoEncoder(output_units=x_train_scaled.shape[1])
model.compile(loss='msle', metrics=['mse'], optimizer='adam')
epochs = 20

history = model.fit(
    x_train_scaled,
    x_train_scaled,
    epochs=epochs,
    batch_size=512,
    validation_data=(x_test_scaled, x_test_scaled)
)
```

```
Epoch 1/20
5/5 [==============================] - 1s 43ms/step - loss: 0.0110 - m
se: 0.0243 - val_loss: 0.0132 - val_mse: 0.0301
Epoch 2/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0106 - m
se: 0.0234 - val_loss: 0.0129 - val_mse: 0.0295
Epoch 3/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0098 - m
se: 0.0215 - val_loss: 0.0125 - val_mse: 0.0286
Epoch 4/20
5/5 [==============================] - 0s 15ms/step - loss: 0.0087 - m
se: 0.0193 - val_loss: 0.0122 - val_mse: 0.0278
Epoch 5/20
5/5 [==============================] - 0s 15ms/step - loss: 0.0077 - m
se: 0.0171 - val_loss: 0.0115 - val_mse: 0.0264
Epoch 6/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0068 - m
se: 0.0151 - val_loss: 0.0110 - val_mse: 0.0252
Epoch 7/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0061 - m
se: 0.0135 - val_loss: 0.0104 - val_mse: 0.0239
Epoch 8/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0056 - m
se: 0.0124 - val_loss: 0.0101 - val_mse: 0.0233
Epoch 9/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0053 - m
se: 0.0118 - val_loss: 0.0100 - val_mse: 0.0230
Epoch 10/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0051 - m
se: 0.0113 - val_loss: 0.0100 - val_mse: 0.0229
Epoch 11/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0049 - m
se: 0.0109 - val_loss: 0.0099 - val_mse: 0.0227
Epoch 12/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0049 - m
se: 0.0108 - val_loss: 0.0098 - val_mse: 0.0226
Epoch 13/20
5/5 [==============================] - 0s 13ms/step - loss: 0.0048 - m
se: 0.0106 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 14/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0047 - m
se: 0.0105 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 15/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0047 - m
se: 0.0104 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 16/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0046 - m
```

```
se: 0.0103 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 17/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0046 - m
se: 0.0102 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 18/20
5/5 [==============================] - 0s 13ms/step - loss: 0.0046 - m
se: 0.0101 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 19/20
5/5 [==============================] - 0s 12ms/step - loss: 0.0045 - m
se: 0.0100 - val_loss: 0.0098 - val_mse: 0.0227
Epoch 20/20
5/5 [==============================] - 0s 11ms/step - loss: 0.0045 - m
se: 0.0099 - val_loss: 0.0098 - val_mse: 0.0226
```

In [15]:

```
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.xlabel('Epochs')
plt.ylabel('MSLE Loss')
plt.legend(['loss', 'val_loss'])
plt.show()
```



Finding threshold for anomaly and doing predictions

In [17]:

```
def find_threshold(model, x_train_scaled):
  reconstructions = model.predict(x_train_scaled)
  reconstruction_errors = tf.keras.losses.msle(reconstructions, x_train_scaled)
  threshold = np.mean(reconstruction_errors.numpy()) \
   + np.std(reconstruction_errors.numpy())
  return threshold

def get_predictions(model, x_test_scaled, threshold):
  predictions = model.predict(x_test_scaled)
  errors = tf.keras.losses.msle(predictions, x_test_scaled)
  anomaly_mask = pd.Series(errors) > threshold
  preds = anomaly_mask.map(lambda x: 0.0 if x == True else 1.0)
  return preds

threshold = find_threshold(model, x_train_scaled)
print(f"Threshold: {threshold}")
```

```
Threshold: 0.009662778406606926
```

Getting accuracy score

In [18]:

```
predictions = get_predictions(model, x_test_scaled, threshold)
accuracy_score(predictions, y_test)
```

Out[18]:

```
0.951
```

In [ ]:

```

```