

1. Explain your implementation which get the best performance in detail.

這次作業中，我實作了 KNN 和 auto encoder 兩種方法。經過比較和測試，發現 auto encoder 的結果比較好，可以在 public testing data 達到高於 0.95 的分數。我的 model 在 encoder 的部分使用了兩個 linear 層，並在中間使用 ReLU 作為 activation function 連接到 decoder，並將 encoder 的 output 用兩個 linear 層穿插 ReLU 產生 decoder 的結果。我使用了每樣資料中提供的 16 種數據做為 input，沒有特別作 feature selection，所以 auto encoder 的 input dimension 就是 16。

```
class Autoencoder(nn.Module):
    def __init__(self, input_dim):
        super(Autoencoder, self).__init__()
        self.encoder = nn.Sequential(
            nn.Linear(input_dim, 8),
            nn.ReLU(),
            nn.Linear(8, 4),
            nn.ReLU()
        )
        self.decoder = nn.Sequential(
            nn.Linear(4, 8),
            nn.ReLU(),
            nn.Linear(8, input_dim)
        )

    def forward(self, x):
        encoded = self.encoder(x)
        decoded = self.decoder(encoded)
        return decoded
```

訓練過程中我分別對每個不同字母的 training data 進行訓練，將每次訓練後的字母模型對 testing data 進行重建，儲存重建後的 reconstruction loss，並在全部計算完後，取所有字母模型計算結果中 loss 最低的那個做為輸出，因為每個模型都是針對一個字母訓練，所以一個正常的字母應該要在對應模型上的 loss 特別低，而 outlier 就會在所有模型的 loss 都偏高。

2. Explain the rationale for using auc score instead of F1 score for binary classification in this homework.

AUC score 先利用 ROC curve 在不同 threshold 下畫出 True Positive 與 False Positive 的曲線，再計算曲線下的面積。AUC 對 data imbalance 不敏感，因為它考慮的是預測的相對好壞而不是絕對數量。而 F1 score 則會因為 data imbalance 而有所偏頗。另外，AUC score 衡量模型在所有可能 threshold 下分類的的能力，能更全面的評估模型性能。F1 score 依賴特定 threshold 而且對 data imbalance 很敏感，相較之下在 anomaly detection 中沒有比 AUC score 更好。

### 3. Discuss the difference between semi-supervised learning and unsupervised learning

Unsupervised learning 只使用 unlabeled data 來發現數據中的模式和結構，而 semi-supervised learning 使用少量標記數據和大量未標記數據，透過結合 labeled 和 unlabeled data 來改善模型。