Report submitted by: Hema

**Purpose of Data Science Project**

The purpose of this data science project is to determine whether we can utilize Bank Authenthication data to differentiate genuine and forged notes and help the bank develop an algorithm to differentiate the two.

**Description of Data**

The data given was available in CSV file from identification of images of banknotes in 2 columns, V1 which is variance of the images and V2 which was the skewness of the images transformed.

**Methods of Data Analysis**

Both the columns were used in the data analysis. First, pandas library was used to load the CSV file and numpy was used to set the data in V1 and V2 into an 2D array. The data was viewed in matplotlib.pyplot for correctness. The mean and standard deviation was calculated, normalized and plotted for both the data, as shown in Figure 1.

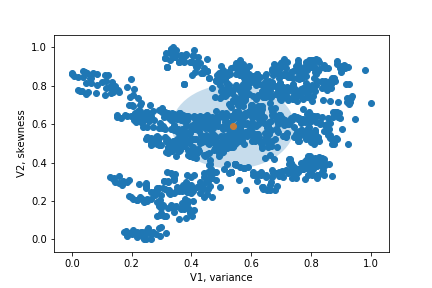


Figure 1

Next, KMeans module was imported was sklearn.cluster and numpy was used to stack both the columns into 2D array. The number of clusters was set to 2 since we want to identify genuine and forged notes and the cluster centers were set to visualize the scatter plot better as shown in Figure 2.

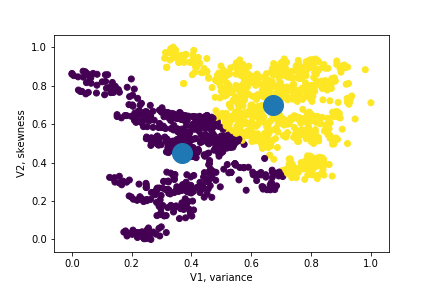


Figure 2

**Results**

The plot showed that there are 2 clusters that are clearly visible and their centroids. After running the model several times, the model detected 765 genuine notes, in purple and 607 fake notes, in yellow as shown in Figure 3. The values of

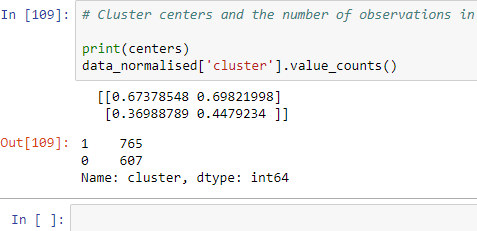


Figure 3

What we can deduce from this scatterplot is that the variance and skewness of the images of the genuine notes in purple are smaller and the data points are less disperse compared to the fake ones in yellow.

After running the model for several times, the number of genuine and forged notes remained the same, suggesting that the model is stable to use.

**Recommendations for Client**

Since the model is stable for use, the client can use this model for the detection of fake notes and the model to be reviewed after some time of test.