ECE 657A Assignment 3
Submitted By:
Amandeep Kaur (21044104)
Bhupesh Dod (21046099)

Question 4: Using Your Own Encoding
**Solution:**

To run the CNNs, we used Google Colab and Keras Library. We used Google Colab GPU to run the CNNs for faster computations. The steps followed to create and implement the own CNN network are as follows:

To run the CNNs, we used Google Colab and Keras Library. The steps followed to create and implement the CNN are as follows:
1. Load all the necessary libraries.
2. Import the Datasets.
3. Data pre-processing. : Reshape the data to 28*28
4. Create compile and summarize the CNN network.
5. Fit the CNN network.
6. Using Model class from Keras library we extract one of the fully connected hidden layers from the CNN model.
7. Apply PCA, DBScan, K-means and T-SNE to the extracted encoded layer and visualize it.

4.1. Visualize your encoding with the first two components from PCA, the colour mapping could be the label values.

Principal Component Analysis with no of components = 2 is performed on the encoder output and the resulting two principal components are plotted (as shown below). The data is segmented into 5 different colour labels depending upon the y_test labels.
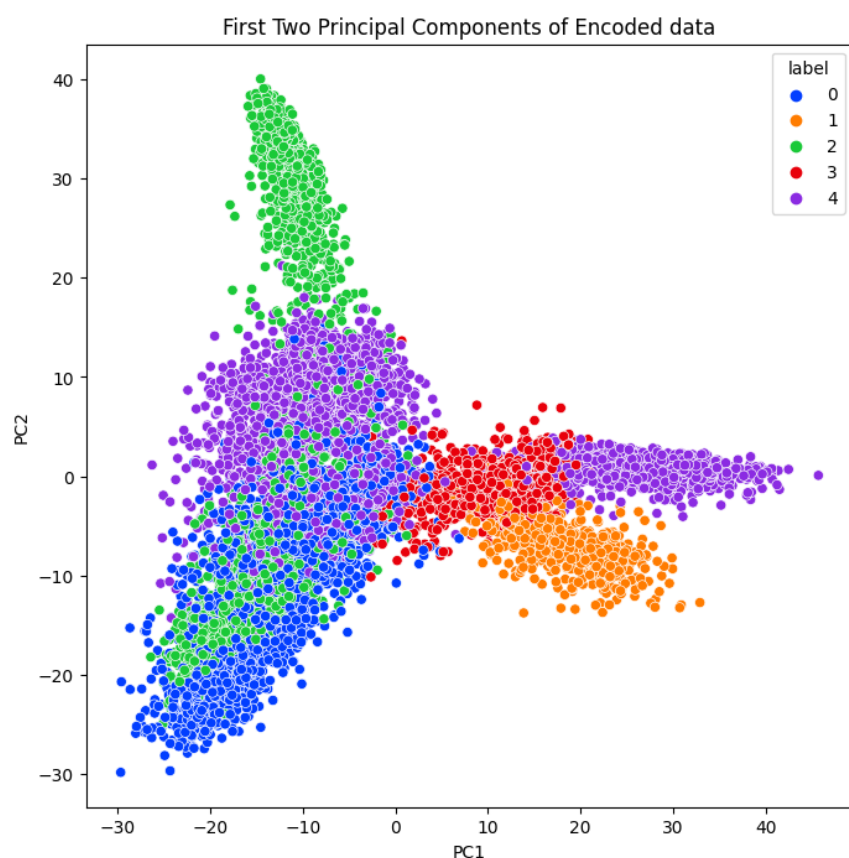

4.2 Applying the DBScan and K-means clustering

Evaluate the DBScan and K-means cluster, then add those labels as colour arguments to the PCA plots. The five clusters can be clearly seen from the plot in both DBScan and K-means, although there was some noise with label = -1, which is removed to compare the two plots.
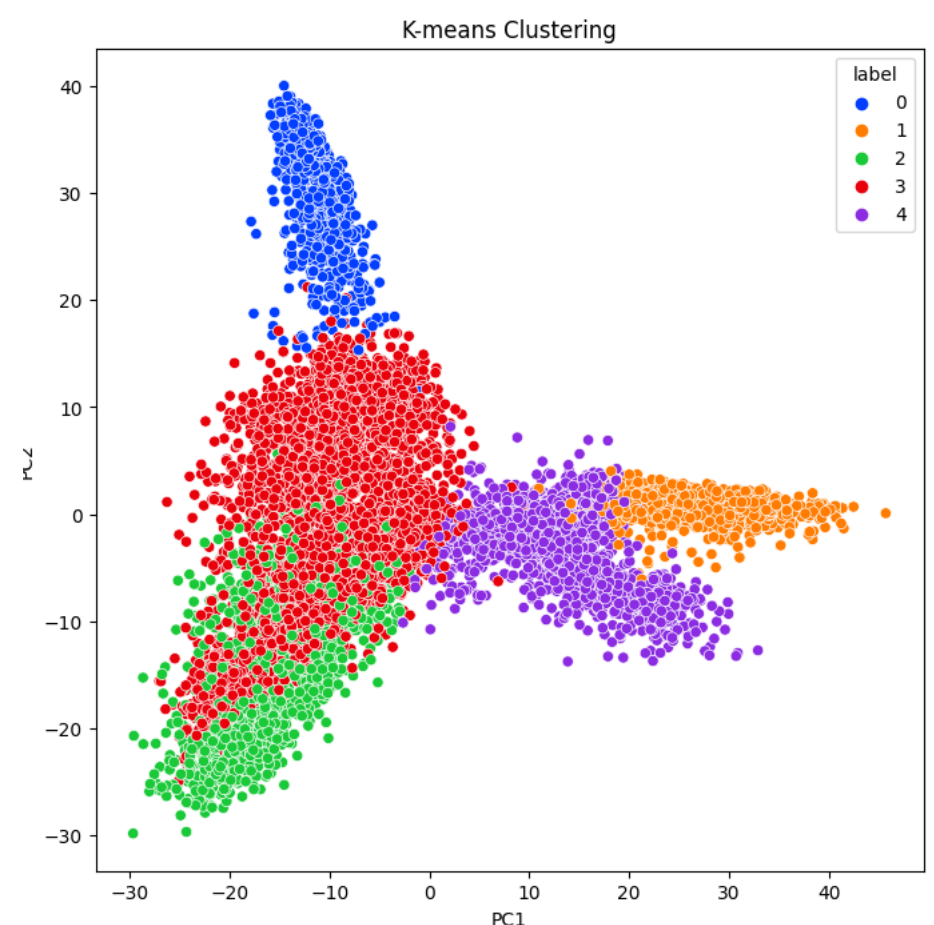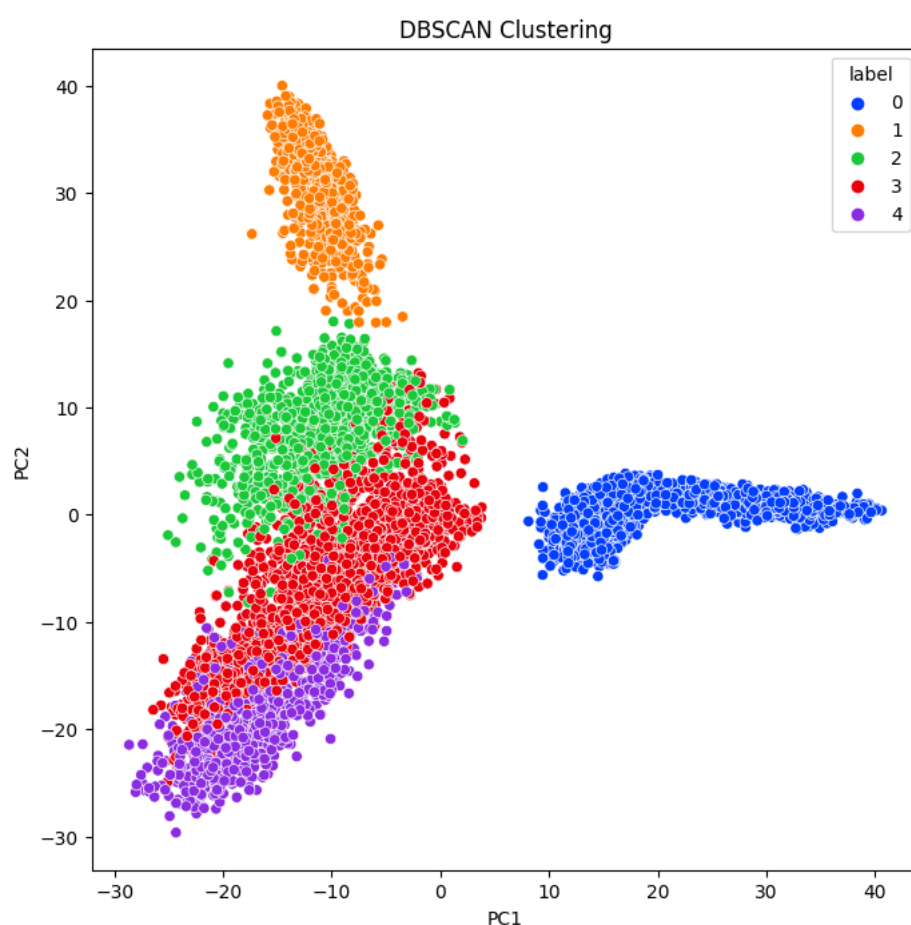
4.3. Apply t-SNE with 2 components on encoder output and plot the result with y_test (test output variable) labels

It can be seen that there are three/two clusters visible with Label 0 which is aligned with the mystery label (T-shirt/Top and Shirt) as we discovered in question-1 of assignment. And all other labels follow the same trend like, only one cluster with Label 1, two clusters with Label 2, two clusters with Label 3, and three clusters with Label 4.
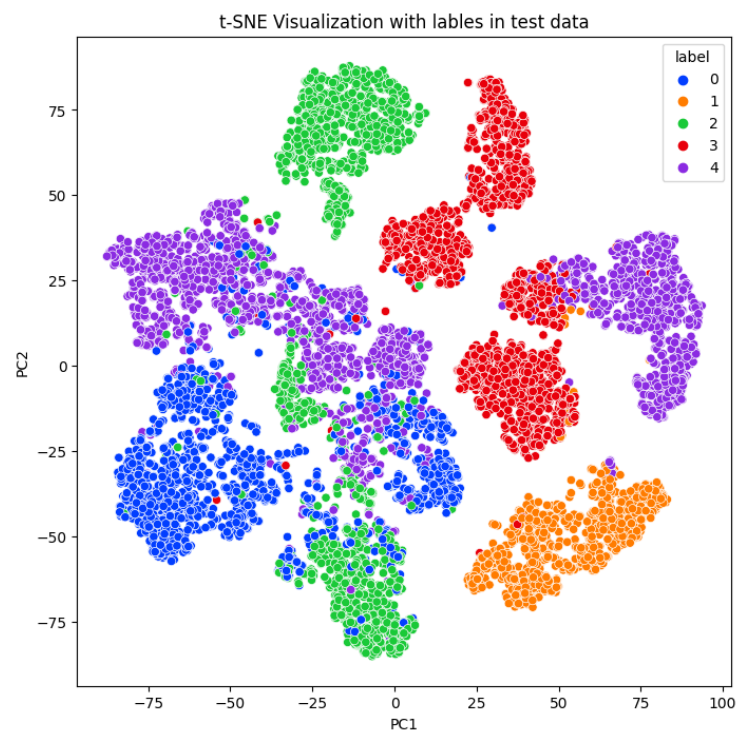
Label	Fashion Item
0	T-shirt/Top and Shirt
1	Sandal
2	Pullover and Trouser
3	Sneaker and Bag
4	Dress, Ankle and Coat

Applying t-SNE with 2 components on encoder output and plotting the result with DBScan and K-means labels.
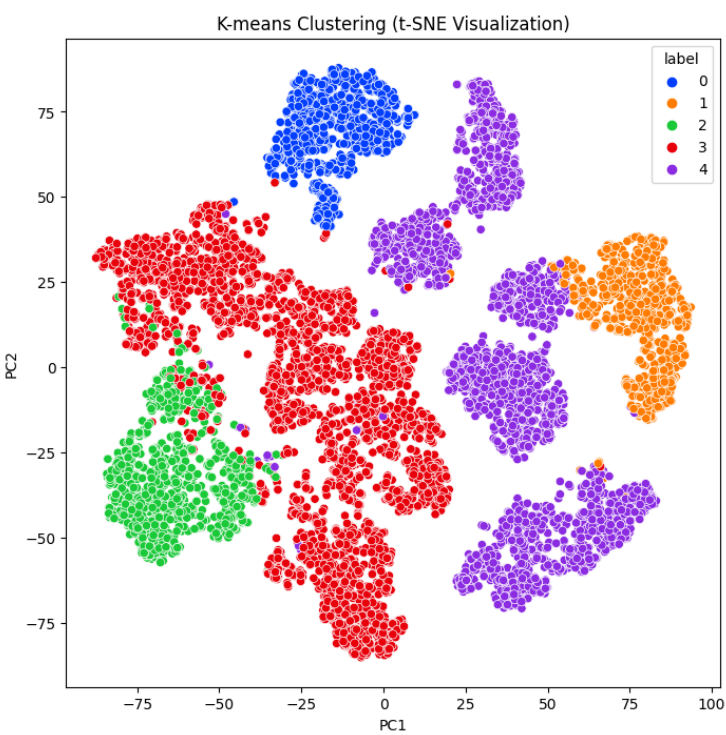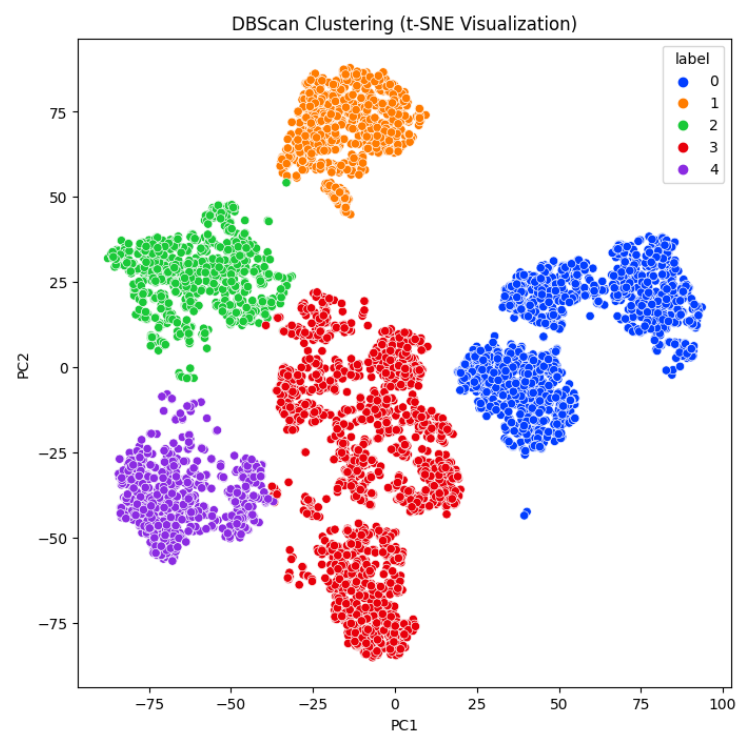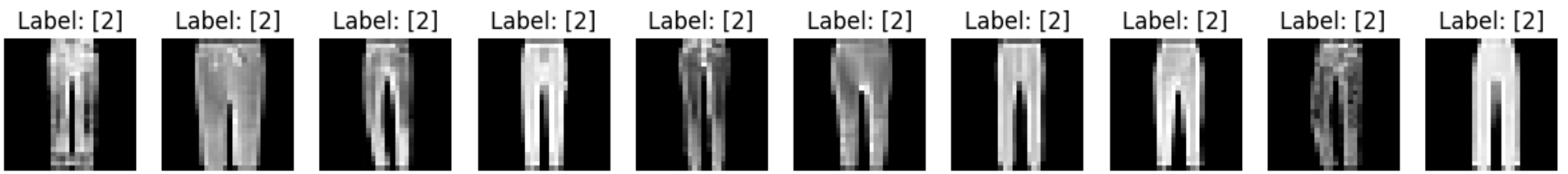

4.4. Guessing the labels

From K-means clustering visualized by t-SNE (above plot), it can be interpreted that there is one cluster with label 0, and one cluster with label 1, one clusters with label 2, three clusters with label 3 and two clusters with label 4.

The new labels assigned by K-means clustering can be interpreted from visualizing the images of the clusters formed by K-Means clustering as shown below.

It can be interpreted as the k-means clusters the labels based on their similarities in the images on datasets, for example, trouser is made a separate cluster as no other item is similar in appearance.

Cluster	Label	Fashion Item
Cluster 0	Label - 2	Trouser
Cluster 1	Label - 4	Ankle Boot
Cluster 2	Label - 0	T-Shirt/Top
Cluster 3	Label – 0, 2, 4	Shirt, Dress, Pullover and Coat
Cluster 4	Label 1, 3	Sandal, Sneaker and Bag

Cluster 0

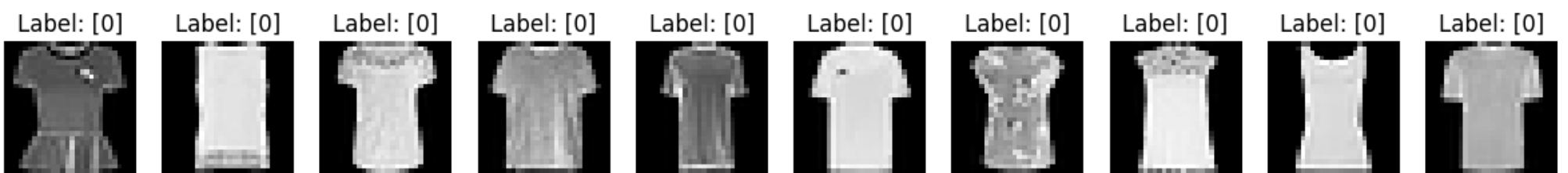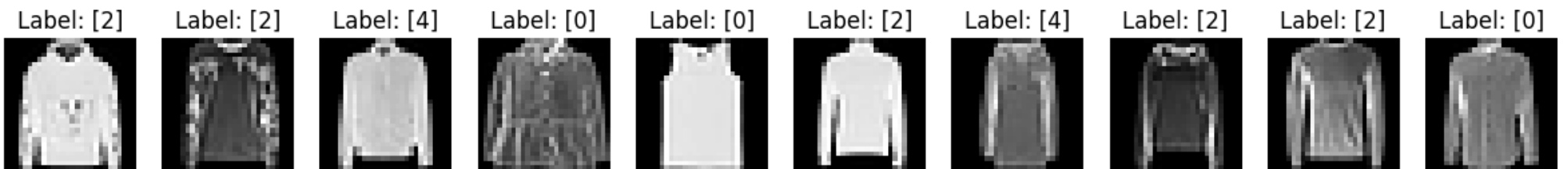| Label: [2]	Label: [2]	Label: [2]	Label: [2]	Label: [2]	Label: [2]	Label: [2]	Label: [2]	Label: [2]	Label: [2]

Cluster 1

Label: [4]	Label: [4]	Label: [4]	Label: [4]	Label: [4]	Label: [4]	Label: [4]	Label: [4]	Label: [4]	Label: [4]

Cluster 2

Label: [0]	Label: [0]	Label: [0]	Label: [0]	Label: [0]	Label: [0]	Label: [0]	Label: [0]	Label: [0]	Label: [0]

Cluster 3

Label: [2]	Label: [2]	Label: [4]	Label: [0]	Label: [0]	Label: [2]	Label: [4]	Label: [2]	Label: [2]	Label: [0]

Cluster 4

Label: [1]	Label: [1]	Label: [3]	Label: [1]	Label: [3]	Label: [1]	Label: [3]	Label: [3]	Label: [3]	Label: [3]