

# Assignment - 2

## Bag of Words Based Categorization for FashionMNIST Dataset

Ashish Sharma  
2019MCB1213

**Note: The code submitted takes about an hour to run to complete**

### 1) Features used:

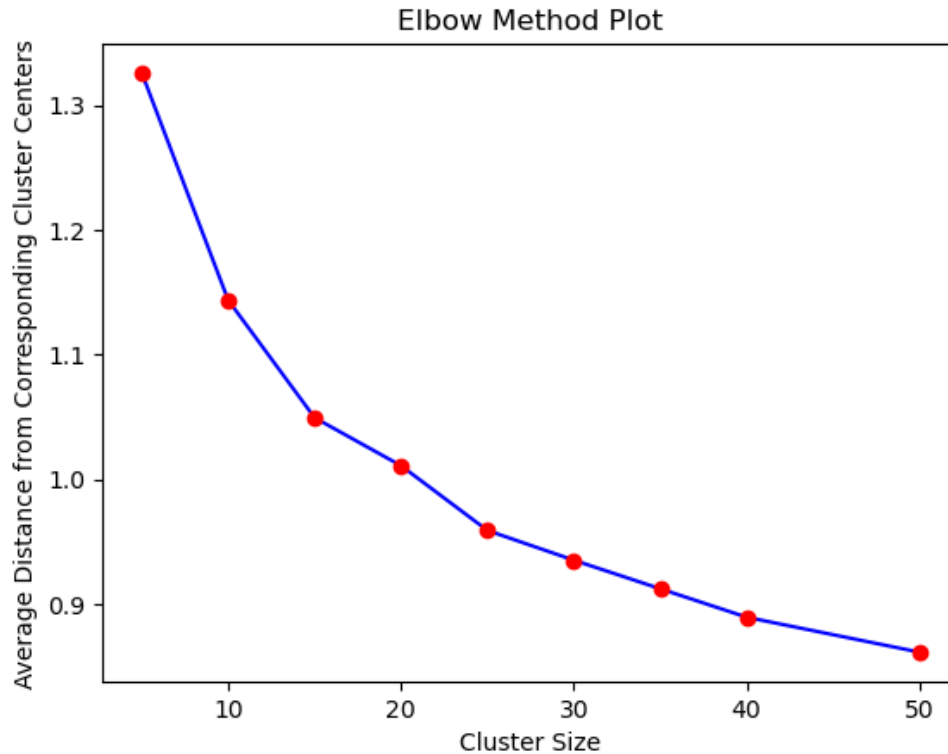
- a) Each image is originally of size 28 x 28, with intensities in the range of 0 to 255.
- a) The pixel intensities were scaled between 0 and 1 by dividing by 255.
- b) Each image was then divided into a 4x4 grid of 16, 7x7 cells of pixels.
- c) **Each 7x7 cell represents a feature** for an image
- d) The 7x7 cells were flattened for computation purposes resulting in feature representation of dimensions, (16, 49), for each image. 16 being number of 7x7 cells, and 49 being the number of pixels inside the 7x7 cell

### 2) How was the number of clusters decided?

- a) Asymptotically it is expected as number of clusters increase the accuracy will also increase on the training data
- b) Too many clusters will cause more computational overhead. Moreover it makes no sense to have too many clusters, as we already have an intuitive upper limit as below for max clusters
- c) A rough estimate of the maximum total number of distinct clusters could be the product of number of classes and number of features for each class, which is  $10 \times 16 = 160$ . Experiment was performed till 50 clusters, for 10 iterations of the k-means algorithm.

- d) The Elbow method was used to decide the number of clusters.  
Experiments were run for various cluster sizes, equally spaced by 5, from 5 through 40, and 50.
- e) **It was observed that the last sharpest drop in average distances occurred for cluster size = 25.**

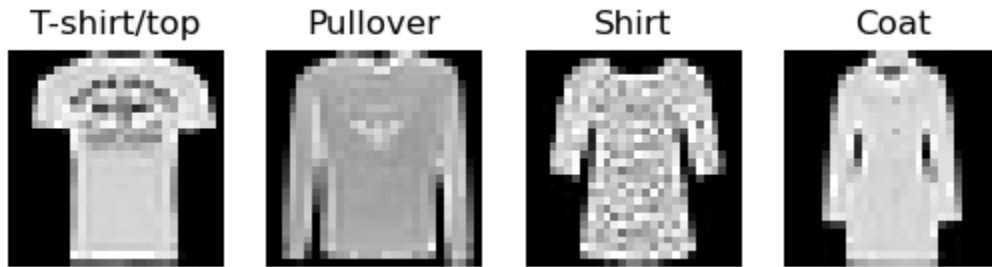
**Hence cluster size of 25 was chosen to train the Bag of Words based Classifier.**



### 3) Experimental Results:

Class Name		precision	recall	f1-score	support	
0	T-shirt/top	0	0.68	0.66	0.67	1000
1	Trouser	1	0.95	0.92	0.93	1000
2	Pullover	2	0.61	0.61	0.61	1000
3	Dress	3	0.76	0.74	0.75	1000
4	Coat	4	0.58	0.60	0.59	1000
5	Sandal	5	0.88	0.84	0.86	1000
6	Shirt	6	0.46	0.50	0.48	1000
7	Sneaker	7	0.85	0.89	0.87	1000
8	Bag	8	0.90	0.85	0.88	1000
9	Ankle boot	9	0.87	0.90	0.89	1000
accuracy				0.75	10000	
macro avg		0.75	0.75	0.75	10000	
weighted avg		0.75	0.75	0.75	10000	

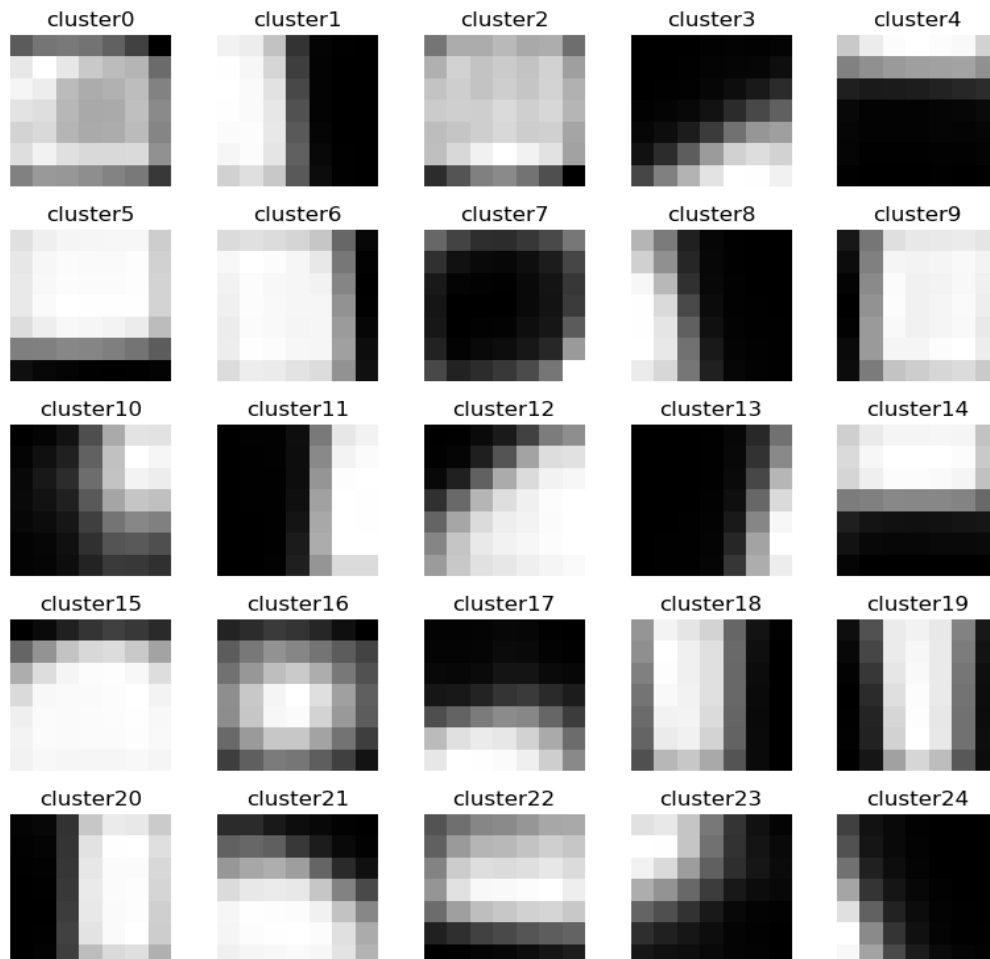
- a) The results obtained on the test set for cluster size of 25 are as above.
- b) The classes for which the recall or the true positive rate are least are T-shirt/top(0), Pullover(2), Coat(4) and Shirt (6)



- c) The overall accuracy obtained was 75% on the test set.

#### 4) Visual Dictionary:

Each **cluster center** is stored in the VisualDictionary directory with each cluster center named as “cluster*i*”, *i* at the end denoting which cluster it represents



## **Appendix: Other Experimental Details**

### **1) Number of Iterations KMeans was run:**

- a) It was observed as the number of iterations increased, although KMeans clustering approached convergence, it performed poorer on the test dataset.
- b) Empirically it was decided to be run on 10 iterations

### **2) Random State Used to Initialize Cluster Centers for the Elbow Plot:**

- a) Sometimes a particular random state would lead to a cluster center getting no points, this would result in a warning from np.mean method.
- b) The outputs too came out to be erroneous at those times
- c) The random state was manually changed and adjusted to avoid whenever such a situation occurred.

### **3) Soft Assignment Method Used in Histogram Calculation:**

- a) Soft assignment was done using Softmax method,  
$$e^{-\alpha * (distance\_of\_current\_cluster\_center - distance\_of\_closest\_cluster\_center)}$$
divided by summation of the above term for all cluster centers for a particular feature. Alpha was empirically chosen to be 2.0