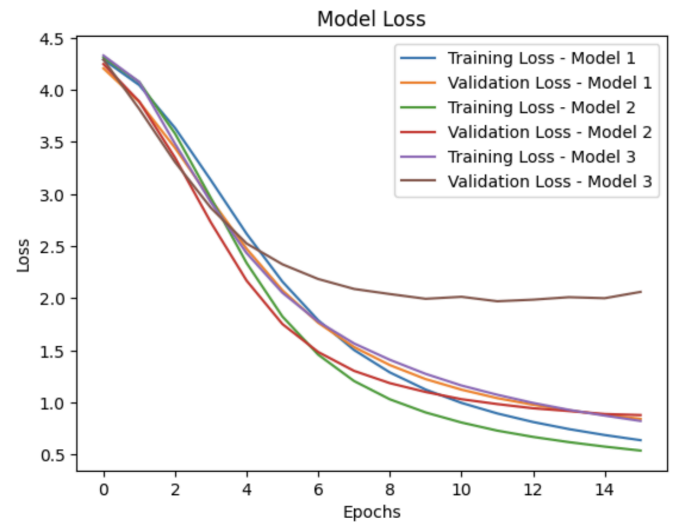
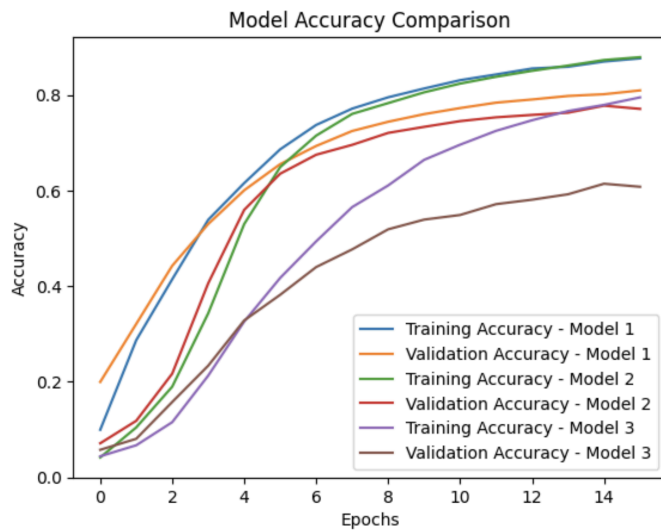


## 1. Neural Network Models on 77 classes:

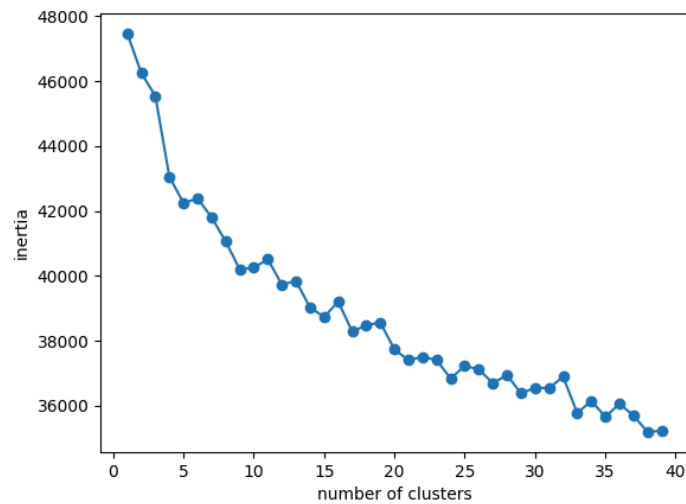
- For the three neural network models (a model with two layers, a model with 3 layers, and a model with 6 layers), their accuracy/loss on the classification on the 77 labels is listed below:

Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Testing Accuracy	Testing Loss
1	0.8695	0.6639	0.8094	0.8443	0.8205	0.8522
2	0.8788	0.5443	0.7707	0.8806	0.7804	0.8927
3	0.7974	0.8264	0.6078	2.0620	0.6054	2.2590



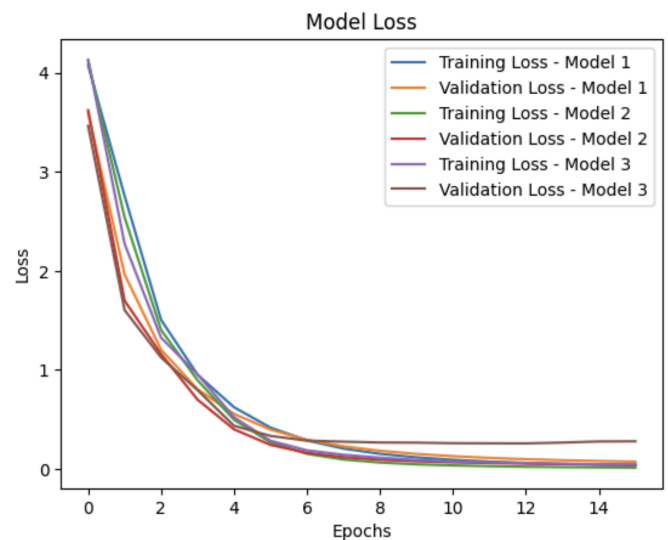
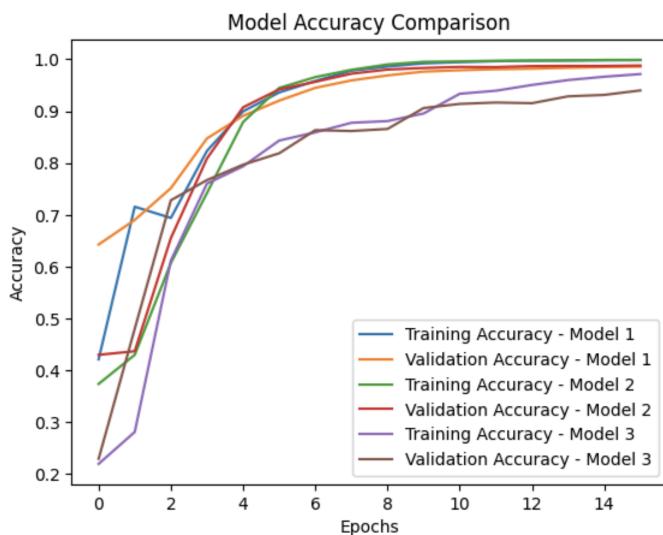
## 2. KMeans Clustering:

- I used the KMeans clustering method to form distinct clusters of the given dataset. I used the elbow method to determine the optimal number of clusters( $k = 8$ ) to use.



- For the three neural network models (a model with two layers, a model with 3 layers, and a model with 6 layers), their accuracy/loss on the classification on the 8 clusters is listed below:

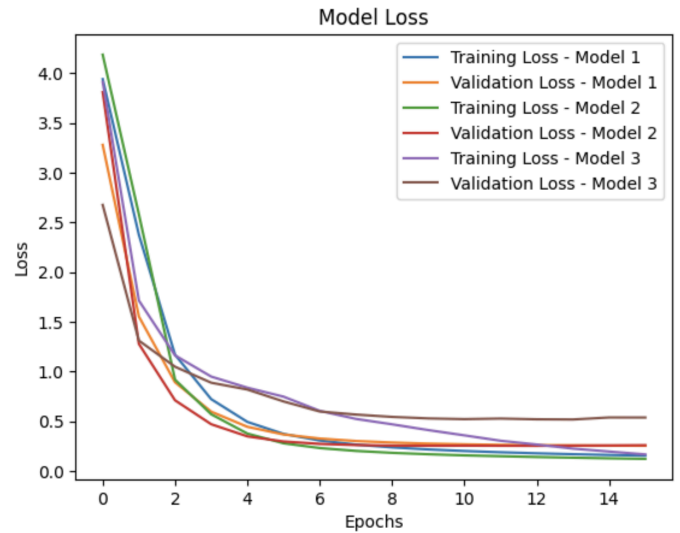
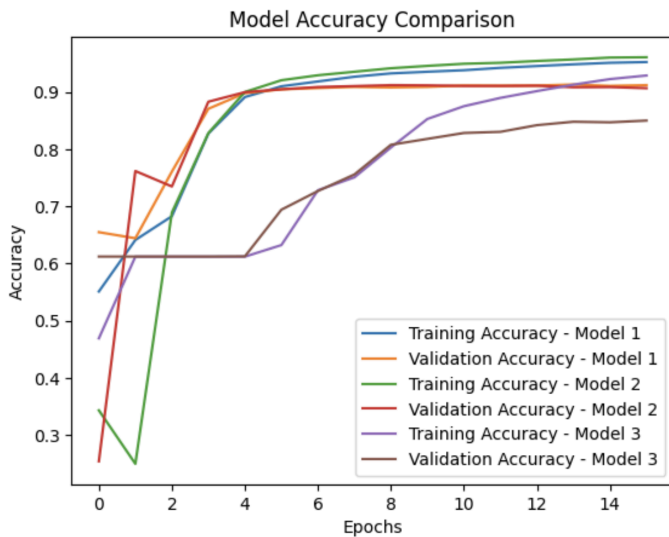
Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Testing Accuracy	Testing Loss
1	0.9994	0.0485	0.9880	0.0793	0.9903	0.0694
2	0.9997	0.0163	0.9850	0.0564	0.9873	0.0514
3	0.9518	0.1530	0.9100	0.3320	0.9363	0.2733



### 3. Hierarchical Clustering:

- I also used the Hierarchical clustering method to form distinct clusters of the given dataset. I used the ward method to cluster the dataset into 4 distinct clusters..
- For the three neural network models (a model with two layers, a model with 3 layers, and a model with 6 layers), their accuracy/loss on the classification on the 4 clusters is listed below:

Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Testing Accuracy	Testing Loss
1	0.9470	0.1604	0.9120	0.2536	0.6678	1.8160
2	0.9621	0.1229	0.9067	0.2580	0.6621	2.3849
3	0.9309	0.1641	0.8500	0.5371	0.7200	1.8598



#### 4. Hierarchical Clustering after dimensionality reduction(PCA):

- After noticing that the models were overfitting and not generalizing well when trained on the clusters formed during hierarchical clustering, I used PCA for dimensionality reduction before clustering. I have used the Hierarchical clustering method to form distinct clusters of the given dataset. I used the ward method to cluster the dataset into 4 distinct clusters..
- For the same neural network model with two layers, its accuracy/loss on the classification on the 4 clusters is listed below:

Metric	Training	Testing
Accuracy	0.9970	0.9913
Loss	0.0083	0.0572
Validation Accuracy	0.9980	N/A
Validation Loss	0.0071	N/A

