CS498 Applied Machine Learning

Homework 3

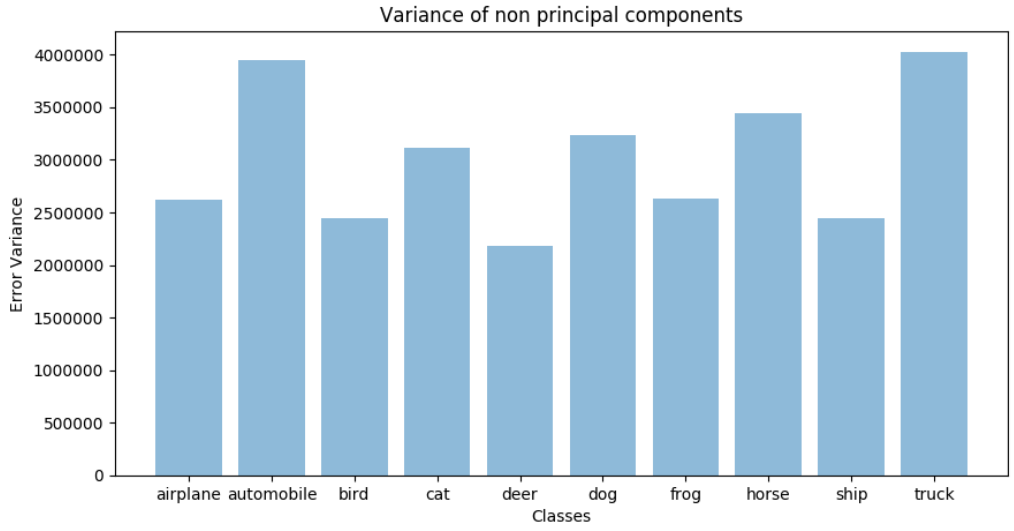
Group Work

Team Leader : Richard Wheeler
NetID: rw6

Team Member: Ling-Hsi Liu
NetID:lhliu2

Part 1:

Plot the error resulting from representing the images of each category using the first 20 principal components

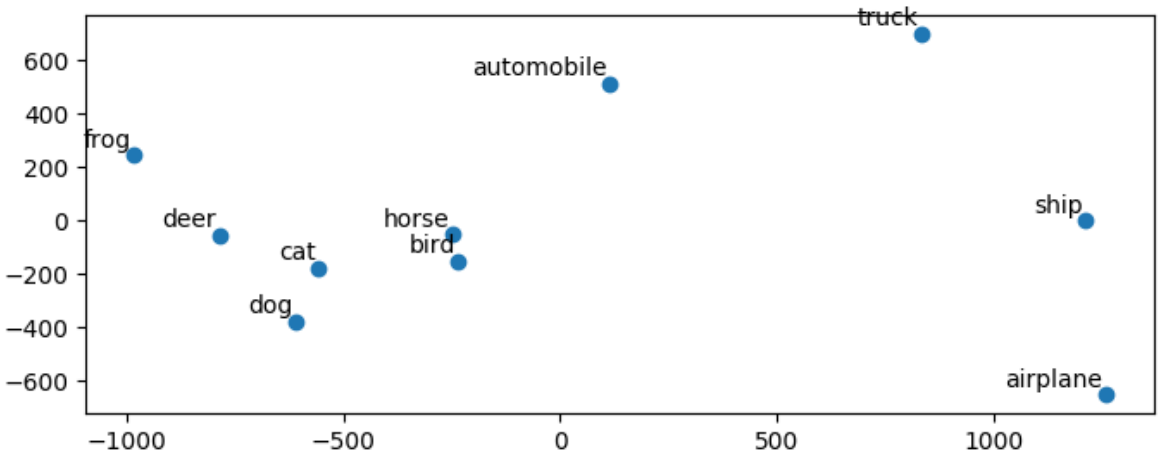


Part 2:

The Distances between mean images for each pair category, 10X10 matrix:

	0	1	2	3	4	5	6	7	8	9
0	0.0	1683.635354...	1605.024349...	1905.535263...	2148.763414...	1965.221491...	2445.679725...	1663.645932...	945.5411038...	1449.0949148...
1	1683.635354...	0.0	886.2367498...	1027.649782...	1143.081367...	1216.079427...	1191.191998...	950.7860778...	1303.466547...	949.99577105...
2	1605.024349...	886.2367498...	0.0	517.3115016...	601.2503352...	701.4682255...	913.7475157...	418.2763060...	1557.715016...	1416.6747332...
3	1905.535263...	1027.649782...	517.3115016...	0.0	469.7917159...	412.1816686...	677.4919693...	596.3767371...	1851.214500...	1676.4678555...
4	2148.763414...	1143.081367...	601.2503352...	469.7917159...	0.0	617.6971411...	460.5109294...	684.3469106...	2065.621656...	1830.7408500...
5	1965.221491...	1216.079427...	701.4682255...	412.1816686...	617.6971411...	0.0	828.5810511...	843.6720912...	1897.591821...	1880.2437717...
6	2445.679725...	1191.191998...	913.7475157...	677.4919693...	460.5109294...	828.5810511...	0.0	948.7040208...	2249.199783...	1913.2408803...
7	1663.645932...	950.7860778...	418.2763060...	596.3767371...	684.3469106...	843.6720912...	948.7040208...	0.0	1660.268079...	1347.3341022...
8	945.5411038...	1303.466547...	1557.715016...	1851.214500...	2065.621656...	1897.591821...	2249.199783...	1660.268079...	0.0	1066.9416321...
9	1449.094914...	949.9957710...	1416.674733...	1676.467855...	1830.740850...	1880.243771...	1913.240880...	1347.334102...	1066.941632...	0.0

Plot: Using principal coordinate analysis to make a 2D map of the means of each category.

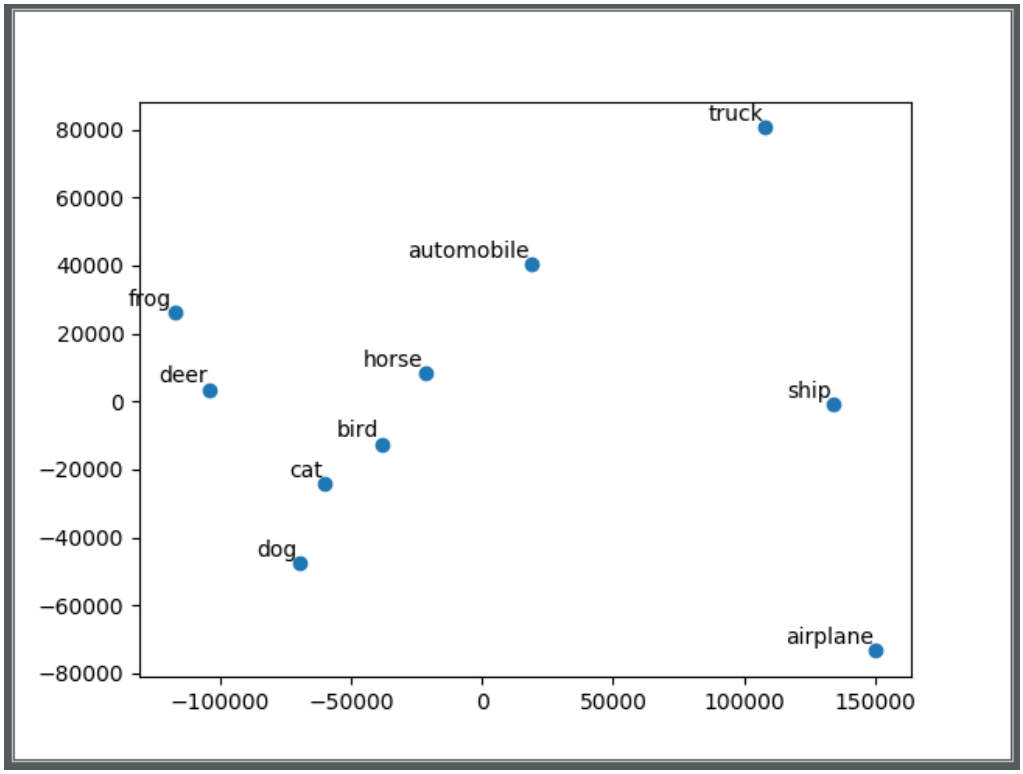


Part 3:

The measure error value for similarity between each category, the distance is ½ (E(A ->B)+E(B->A)) Matrix 10X10 for principal coordinate analysis

	0	1	2	3	4	5	6	7	8	9
0	0.0	181703.9567355896	201600.13644169655	219044.73493773845	268532.523709035	231323.39373264127	288652.64441220515	194729.4813383185	114023.1233845472	162530.08127355707
1	181703.95673558928	0.0	94869.82353554928	111525.19105884898	137519.1513733857	134276.3208092113	142439.59921332484	90820.87399016687	137792.57229154606	115478.76722550181
2	201600.13644169687	94869.82353554928	0.0	51198.70753196774	72184.31480392086	77749.7959524875	95057.14128615467	52318.99666210515	184758.28267879158	178897.22737333152
3	219044.73493773845	111525.19105884898	51198.70753196774	0.0	70121.9781983262	52161.54173905536	83590.90975295642	69934.12481061628	209026.81795313733	200042.8196883836
4	268532.523709035	137519.1513733857	72184.31480392086	70121.9781983262	0.0	82807.01635092661	39864.468344778725	95430.64068824513	244924.10880355287	230469.0877438018
5	231323.39373264127	134276.3208092113	77749.7959524875	52161.54173905536	82807.01635092661	0.0	99226.2813899094	100309.35564691397	215456.8567370519	225101.66487436774
6	288652.64441220515	142439.59921332484	95057.14128615467	83590.90975295642	39864.468344778725	99226.2813899094	0.0	111549.80555140012	258985.78558133298	235591.18282660935
7	194729.4813383185	90820.87399016687	52318.99666210515	69934.12481061628	95430.64068824513	100309.35564691397	111549.80555140012	0.0	178766.6261039661	152933.67367069633
8	114023.1233845472	137792.5722915465	184758.28267879126	209026.81795313733	244924.10880355287	215456.8567370519	258985.78558133298	178766.6261039661	0.0	117662.9862355851
9	162530.08127355707	115478.76722550234	178897.22737333117	200042.8196883836	230469.0877438018	225101.66487436774	235591.18282660935	152933.67367069633	117662.9862355851	0.0

Plot: 2D plot by the similarity distance between each category



In turn, the similarity measure should be big. Use principal coordinate analysis to make a 2D map of the classes. Compare this map to the map in the previous exercise? Are they different? why?

The 2D maps in Part 2 and Part 3 are very different because they show different information. In the part 2, it shows the relative similarity between each mean image of each category. For instance, the mean image of horse (category 7) and bird (category 2) are similar, the value is small. However, comparing the mean image truck (category 9) and dog (category 5), the value is very big, so it is big distance between two categories. In this map from part2, we could see the mean image of truck and airplane has big distances with other categories. And, the value in these two groups is big.

In the part 3, it shows the measure of the similarity of two categories, and we use 20 principal components from each category. In this part, the image of cat and bird are very similar, and the error value from cat (category 3) and bird (category 2) is very small. Otherwise, comparing airplane (category 0) and frog (category 6), the error value is very big. And, when we see the map, the distance is very big between airplane and frog.