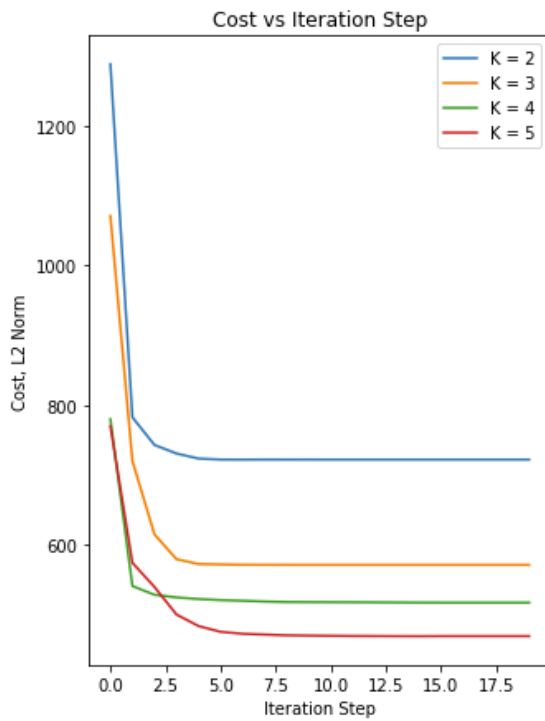
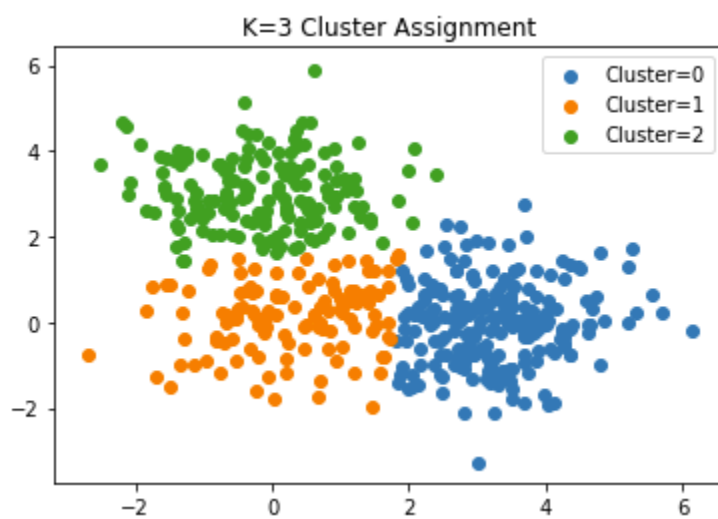


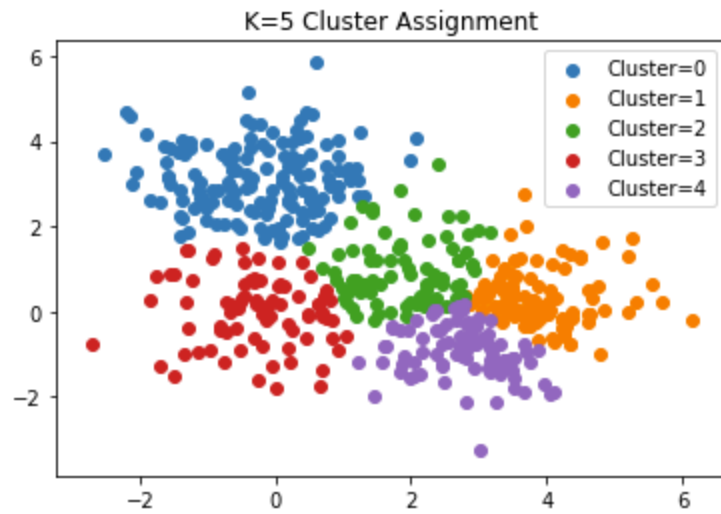
Problem 1.a

First 20 Iterations for $k=2..5$



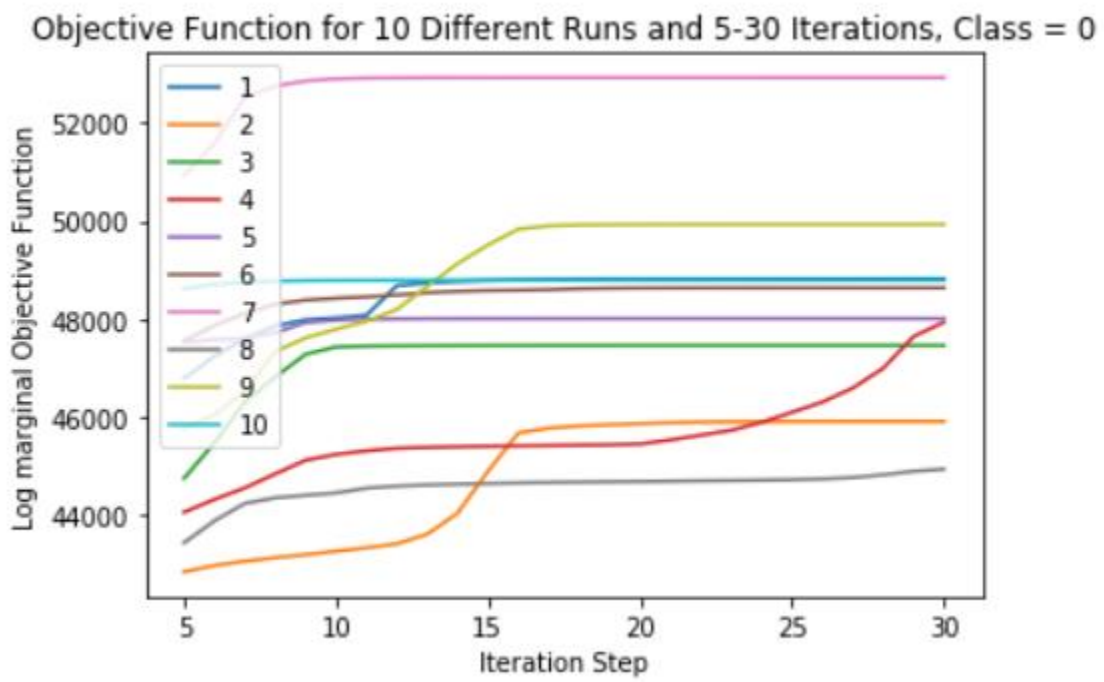
Problem 1.b

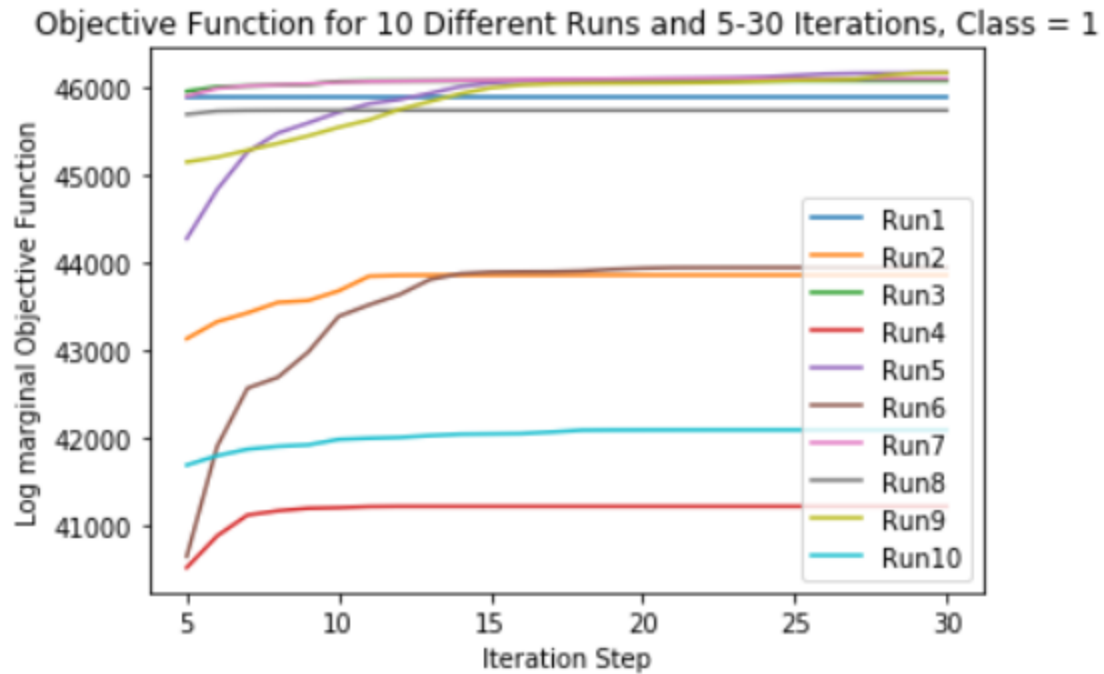




Problem 2.a

EM Algorithm Objective Function Plots





Problem 2.b

Confusion Matrix and Accuracy Using the Best Outputs from EM

	Actual =0	Actual = 1
Prediction = 0	194	7
Prediction = 1	84	175

Accuracy = 80.21%

```
[194, 7]
[84, 175]
accuracy: 0.8021739130434783
```

Repeat for k= 1..4 Accuracies nearby each other. Details and confusion matrix below

K=1 : 77.4%

K=2 : 80%

K=3 : = 81.5%

K=4 : = 82.6%

Confusion matrix and details

K = 1

	Actual =0	Actual = 1
Prediction = 0	183	9
Prediction = 1	95	173

Accuracy = 77.4%

K = 2

	Actual =0	Actual = 1
Prediction = 0	192	6
Prediction = 1	86	176

Accuracy = 80%

K = 3

	Actual =0	Actual = 1
Prediction = 0	205	12
Prediction = 1	73	170

Accuracy = 81.5%

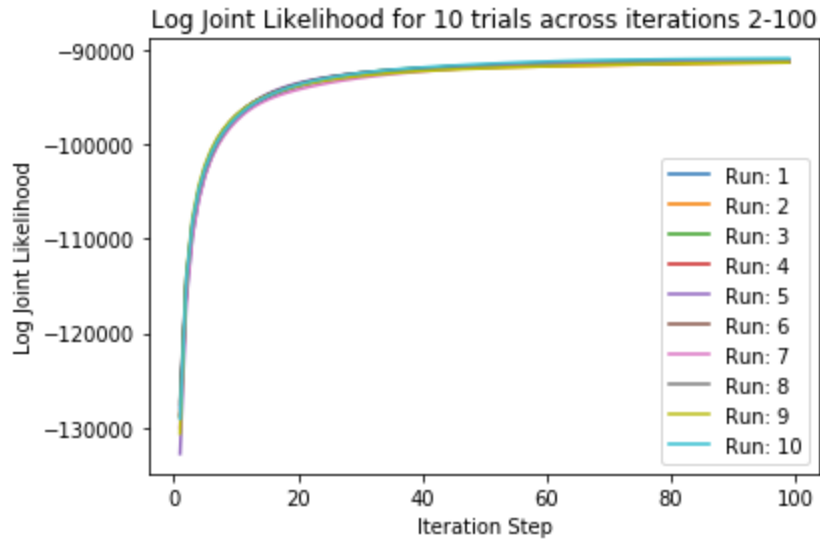
K = 4

	Actual =0	Actual = 1
Prediction = 0	210	12
Prediction = 1	68	170

Accuracy = 82.6%

Problem 3.a

Joint log likelihood of Iterations 2 to 100 for 10 different runs



	joint_log_likelihood	test_set_RMSE
10	-90934.018344	1.130670
1	-91065.581587	1.104772
7	-91177.001490	1.131800
3	-91220.628476	1.107440
6	-91227.346308	1.130560
5	-91231.434833	1.111096
4	-91276.418550	1.149899
8	-91297.915979	1.128656
9	-91353.129632	1.092879
2	-91361.308203	1.147206

Trial 10 has the highest final joint log likelihood but not the lowest RMSE.

Problem 3.b

Below are lists of Starwars, my fair lady, and good fellas movies in their respective orders and their 10 closest movies and its distances sorted by closest movies.

TWO COLUMNS(MOVIES AND DISTANCES) EACH IN FOLLOWING ORDER AND SORTED BY CLOSEST MOVIES:

STAR WARS, MY FAIR LADY, GOOD FELLAS

	star_wars_nearest_movies	star_wars_distances	my_fair_lady_nearest_movies	my_fair_lady_distances	good_fellas_nearest_movies	good_fellas_distances
1	Empire Strikes Back, The (1980)\n	0.293517	My Family (1995)\n	0.779790	Casino (1995)\n	0.578276
2	Raiders of the Lost Ark (1981)\n	0.559219	Mary Poppins (1964)\n	0.820708	Apocalypse Now (1979)\n	0.647309
3	Return of the Jedi (1983)\n	0.601960	Victor/Victoria (1982)\n	0.898342	Bonnie and Clyde (1967)\n	0.660214
4	Day the Earth Stood Still, The (1951)\n	0.679226	SubUrbia (1997)\n	0.905649	Godfather, The (1972)\n	0.771200
5	Usual Suspects, The (1995)\n	0.792622	American in Paris, An (1951)\n	0.937764	Full Metal Jacket (1987)\n	0.775229
6	Indiana Jones and the Last Crusade (1989)\n	0.885387	Winnie the Pooh and the Blustery Day (1968)\n	0.954901	People vs. Larry Flynt, The (1996)\n	0.776257
7	Love in the Afternoon (1957)\n	0.893234	Return of Martin Guerre, The (Retour de Martin...\n	0.992021	Good, The Bad and The Ugly, The (1966)\n	0.777804
8	Blues Brothers, The (1980)\n	0.912179	Sabrina (1954)\n	0.994884	Godfather: Part II, The (1974)\n	0.820151
9	Terminator, The (1984)\n	0.925282	Niçaznette et Boni (1996)\n	1.005537	Once Upon a Time in the West (1969)\n	0.849363
10	Princess Bride, The (1987)\n	0.926964	Cinderella (1950)\n	1.018067	Carlito's Way (1993)\n	0.919316