INF-2B Natural Image Classification Task 3 Report

Task 4.1:

K-NN classification with 100 feature:

		Predicted Class									
		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8	Class9	Class10
Actual Class	Class1	83	0	2	4	1	2	5	0	1	2
	Class2	2	84	2	3	0	1	4	1	2	1
	Class3	0	2	85	1	3	3	1	1	2	2
	Class4	2	2	1	92	1	0	1	1	0	0
	Class5	0	1	4	0	84	6	0	3	2	0
	Class6	0	0	3	1	0	83	0	5	1	7
	Class7	3	2	2	1	1	1	89	1	0	0
	Class8	0	1	5	0	4	10	0	75	3	2
	Class9	0	3	4	3	2	4	0	2	78	4
	Class10	3	0	1	1	0	2	0	0	2	91

Accuracy: 84.4%

K-NN classification with 2 feature:

		Predicted Class									
		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8	Class9	Class10
Actual Class	Class1	41	16	1	17	0	0	14	7	0	4
	Class2	11	28	14	9	0	0	22	3	11	2
	Class3	1	8	42	6	3	0	13	5	22	0
	Class4	28	5	2	21	1	2	5	19	5	12
	Class5	0	0	3	3	68	3	2	9	12	0
	Class6	0	2	1	6	2	68	0	15	2	4
	Class7	15	24	14	10	0	0	24	2	8	3
	Class8	6	3	8	7	4	18	4	26	6	18
	Class9	5	7	13	5	7	2	8	9	39	5
	Class10	7	1	0	16	1	9	1	18	2	45

Accuracy: 40.2%

Gaussian full classification with 100 feature:

		Predicted Class									
		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8	Class9	Class10
Actual Class	Class1	92	0	0	5	0	0	3	0	0	0
	Class2	4	84	4	2	0	1	4	1	0	0
	Class3	2	3	78	2	3	0	6	2	4	0
	Class4	2	0	0	93	2	0	2	0	0	1
	Class5	0	0	1	1	88	4	0	1	4	1
	Class6	2	0	4	2	0	86	1	3	1	1
	Class7	2	1	0	1	0	0	95	1	0	0
	Class8	1	0	3	1	1	9	1	81	2	1
	Class9	0	2	2	2	1	0	1	0	88	4
	Class10	4	1	1	0	1	2	1	0	2	88

Accuracy: 87.3%

Gaussian full classification with 2 feature:

		Predicted Class									
		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8	Class9	Class10
Actual Class	Class1	46	8	0	31	0	0	7	6	0	2
	Class2	8	28	6	5	0	0	32	4	17	0
	Class3	0	0	49	8	0	0	7	4	32	0
	Class4	14	0	0	58	0	0	4	10	7	7
	Class5	0	0	0	0	69	1	0	10	20	0
	Class6	0	0	1	3	0	71	0	18	1	6
	Class7	17	20	8	14	0	0	26	2	12	1
	Class8	3	0	1	10	3	7	2	54	9	11
	Class9	0	0	9	13	3	0	3	10	62	0
	Class10	1	0	1	21	0	2	1	17	1	56

Accuracy: 51.9%

Gaussian lda classification with 100 feature:

		Predicted Class									
		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8	Class9	Class10
Actual Class	Class1	85	0	1	2	2	1	8	0	0	1
	Class2	0	82	4	3	0	1	2	5	2	1
	Class3	1	3	85	0	1	2	1	5	2	0
	Class4	3	0	2	88	2	0	1	2	0	2
	Class5	0	0	8	0	78	4	0	8	2	0
	Class6	1	0	2	2	0	86	0	6	2	1
	Class7	1	4	2	1	0	1	88	2	0	1
	Class8	0	0	6	0	0	2	1	88	2	1
	Class9	0	2	3	2	1	1	0	3	85	3
	Class10	2	0	2	0	0	4	0	1	3	88

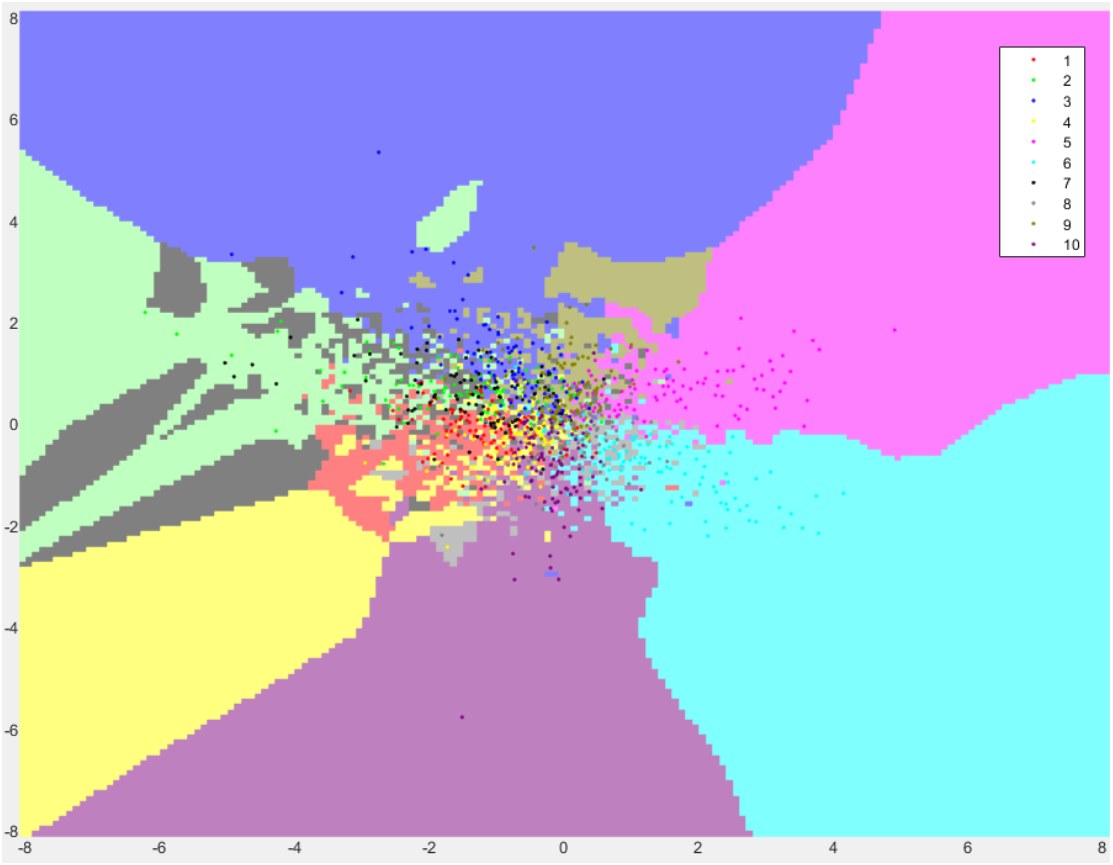
Accuracy: 85.3%

Gaussian lda classification with 2 feature:

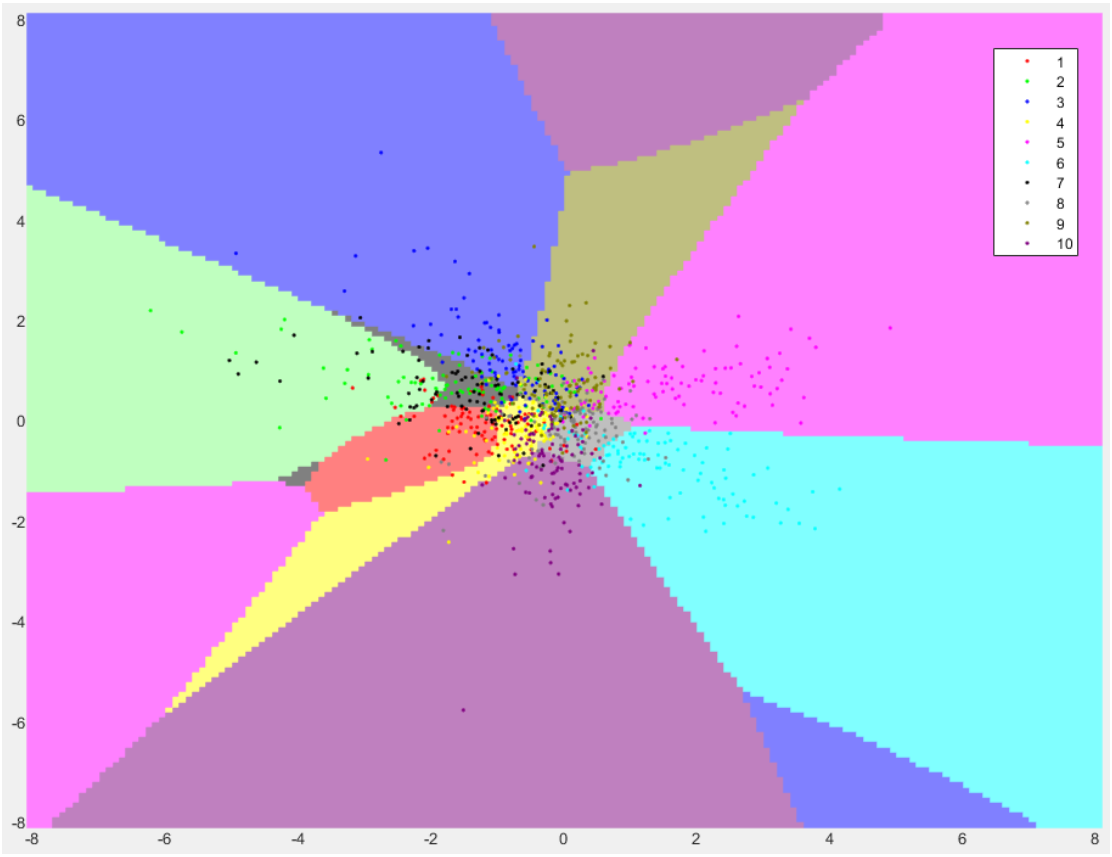
		Predicted Class									
		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8	Class9	Class10
Actual Class	Class1	47	10	0	19	0	0	9	7	0	8
	Class2	7	38	8	3	0	0	23	5	16	0
	Class3	0	2	52	3	0	0	6	7	30	0
	Class4	16	0	1	35	0	0	5	20	6	17
	Class5	0	0	0	0	65	1	0	7	27	0
	Class6	0	0	1	2	1	70	0	16	1	9
	Class7	13	30	8	10	0	0	20	5	11	3
	Class8	3	1	1	4	3	7	1	49	14	17
	Class9	0	1	16	4	6	0	3	18	51	1
	Class10	0	0	1	16	0	2	1	19	1	60

Accuracy: 48.7%

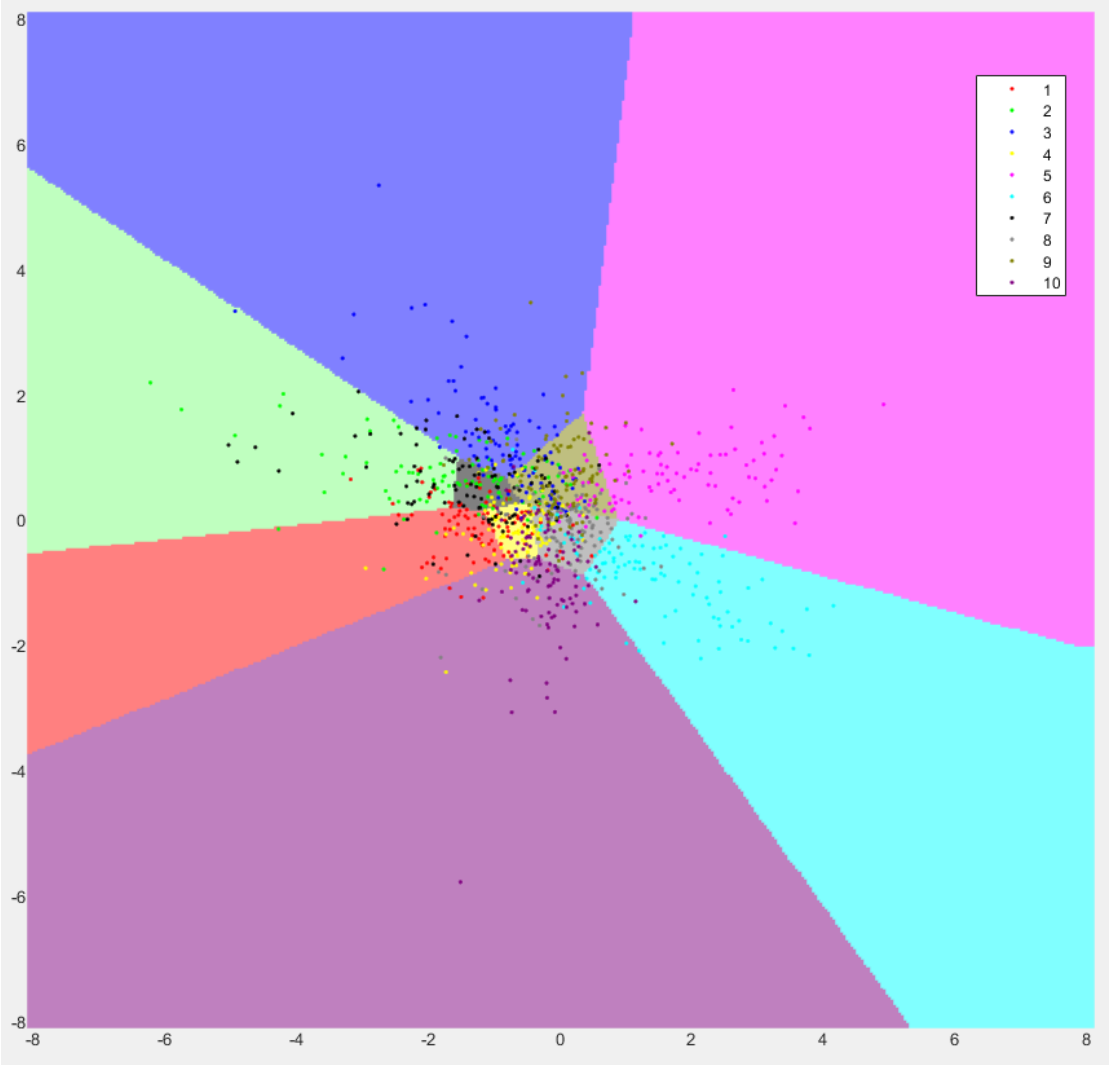
Task 4.2: KNN decision boundary.



Task 4.3: Gaussian-Full decision boundary:



Gaussian lda decision boundaries:



Task 4.4:

From the decision boundary graphs, we can see that the k-nn classification doesn't generalize the classes from the tests, making it having a 'messy' boundary, while both Gaussian method gives clear boundary.

The full Gaussian boundary might be inaccurate when the feature vector is a lot larger than the training data, because there are areas with no data scattered in it.

And the lda boundary have straight line boundaries, it solves the problem from full Gaussian, but makes it less accurate at the center.

From the graph, we can find out that class 2 and 7 (green and black) are not clearly separated, which is also shown in the confusion table, class pairs having the same problem are: 1 and 4(red and yellow), 3 and 9(blue and dark green), this may be the reason that some numbers looks alike when the image isn't that clear.