

INF-2B Natural Image Classification Task 2 Report

Task 3.1:

		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%

Task 3.2:

Determinant of covariance matrix	Class1	4.60989595796651e-234	Class6	1.89603958244386e-241
	Class2	4.59278111007410e-270	Class7	3.04673214719501e-258
	Class3	1.30689372954090e-237	Class8	1.89330949333798e-238
	Class4	1.86676437981820e-238	Class9	2.19296020773058e-244
	Class5	4.72930285263208e-257	Class10	1.53042491859595e-253

		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%

Task 3.3:

Determinant of full covariance matrix: 3.33794166858135e-169

		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%

Task 3.4:

For knn-classification, it has the lowest classification accuracy, and the accuracy even decreases with higher k values, and the decision boundaries aren't clear for low value of k.

The advantage is just easy to implement.

The disadvantage is huge on the other hand, it just compares the test with the train data, not only it's not useful to distinguish features from similar classes, it is very slow when the train data is huge

The full-covariance and lda Gaussian method have the highest classification accuracy, with the longer runtime.

Advantage is that it have the highest accuracy when classifying data like this, giving less mistakes on similar classes, it uses the covariance matrix instead of whole data, making it more efficient.

But due to the large matrix multiplication, it makes the classification slow on high dimensions.