CS669-Pattern Recognition

Assignment - 1

Bayes Classifier

Group 6

Team :

Akhilesh Devrari (B16005)

Aman Jain (B16044)

Yash Agrawal (B16120)

Table of Contents :

1. Objective :

1.1. Build Bayes classifier for classification of following datasets into three classes :

1.1.1. Artificial datasets

1.1.1.1. Linearly separable dataset

1.1.1.2. Non-linearly separable dataset

1.1.2. Real world dataset

1.2. Find confusion matrix for each case (for different datasets).

1.3. Calculate classification accuracy, precision for every class, mean precision, recall for every class, mean recall, F-measure for every class and mean F-measure.

1.4. Plot density contour for each class with data points superposed.

1.5. Plot decision region for each pair of classes and together for all classes.

2. Procedure :

2.1. Separate given data files of each class into training and test data files. Training data consists of 75% and test data consists of 25% of given data set.

2.2. To apply Bayesian classification, assume that data of all classes -follow Gaussian distribution.

2.3. Plot each training data point and density contour for each class.

2.4. Calculate discriminant function (g(x)) for each of the following cases :

- 2.4.1. Covariance matrix for all the classes is the same and is $\sigma^2 I$.
- 2.4.2. Full Covariance matrix for all the classes is the same and is Σ.
- 2.4.3. Covariance matrix is diagonal and is different for each class.
- 2.4.4. Full Covariance matrix for each class is different.

2.5. For each point in X-Y plane calculate its discriminant function for each class and color that point according to the class with maximum value of discriminant function.

2.6. Use test data (25% of total class data) to test the classifier for each case and calculate the confusion matrix, accuracy, recall, precision and F-measure.-

3. Observations :

3.1. Dataset I : Linearly Separable Data

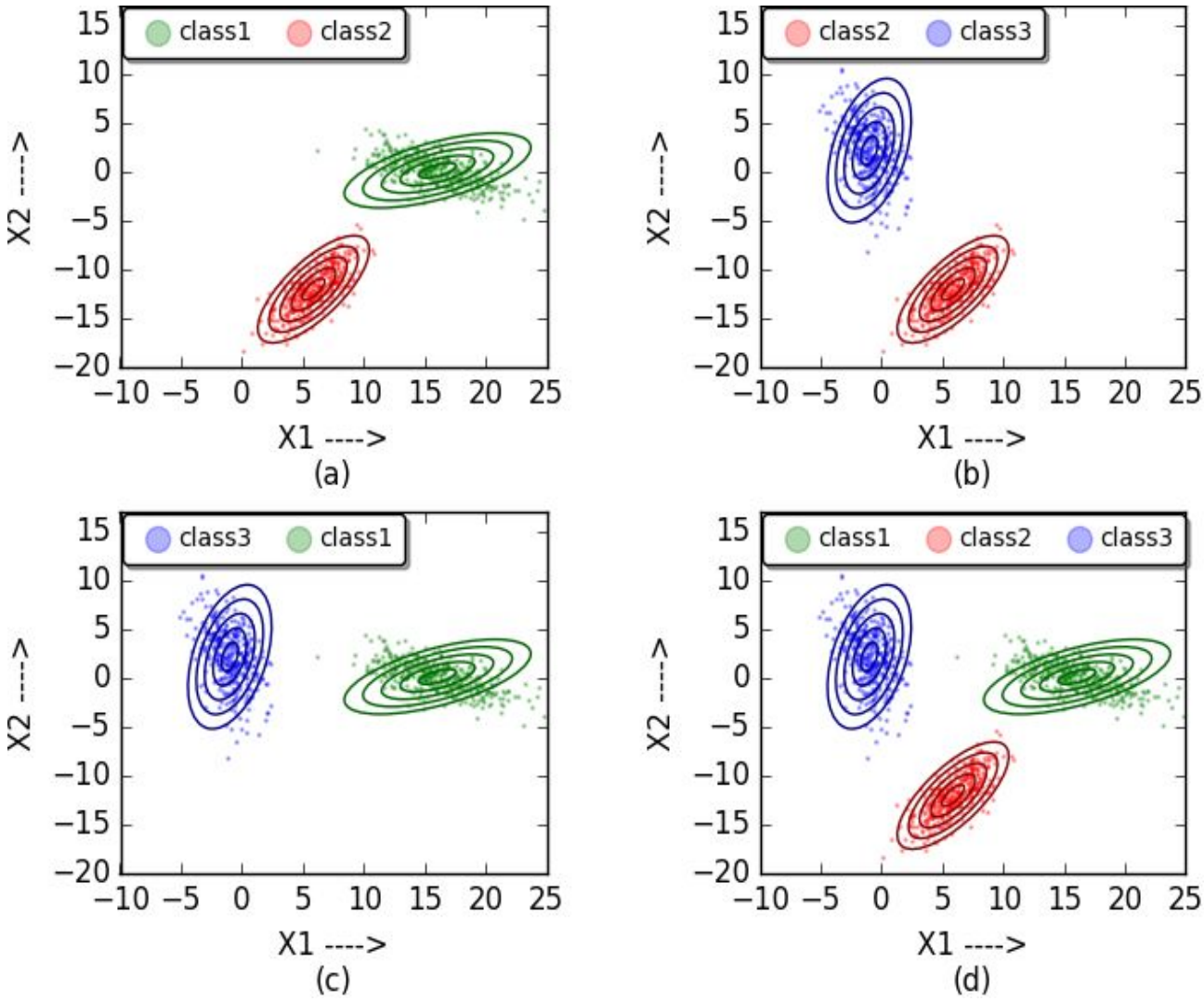

Fig 1 : Contour plot for linearly separable dataset. Contour plot between : (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together

a) Case 1 : Covariance matrix for all the classes is the same and is $\sigma^2 I$

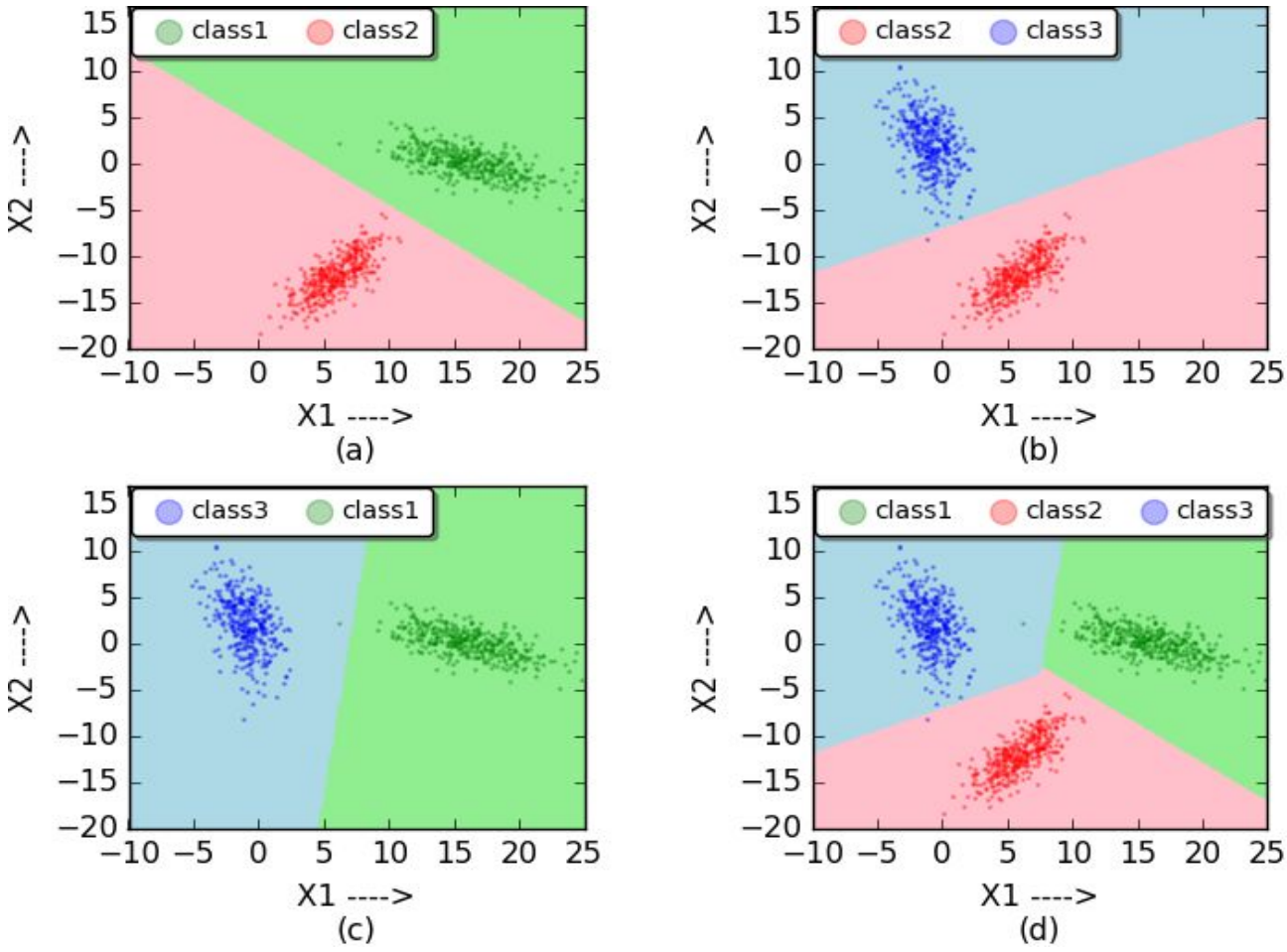

Fig 2 : Class decision boundary plot for case 1, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 100 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	125	0	0
Class 2	0	125	0
Class 3	0	0	125

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	1.0000	1.0000	1.0000	1.0000
Precision	1.0000	1.0000	1.0000	1.0000
F-Measure	1.0000	1.0000	1.0000	1.0000

b) Case 2 : Full Covariance matrix for all classes is the same and is Σ

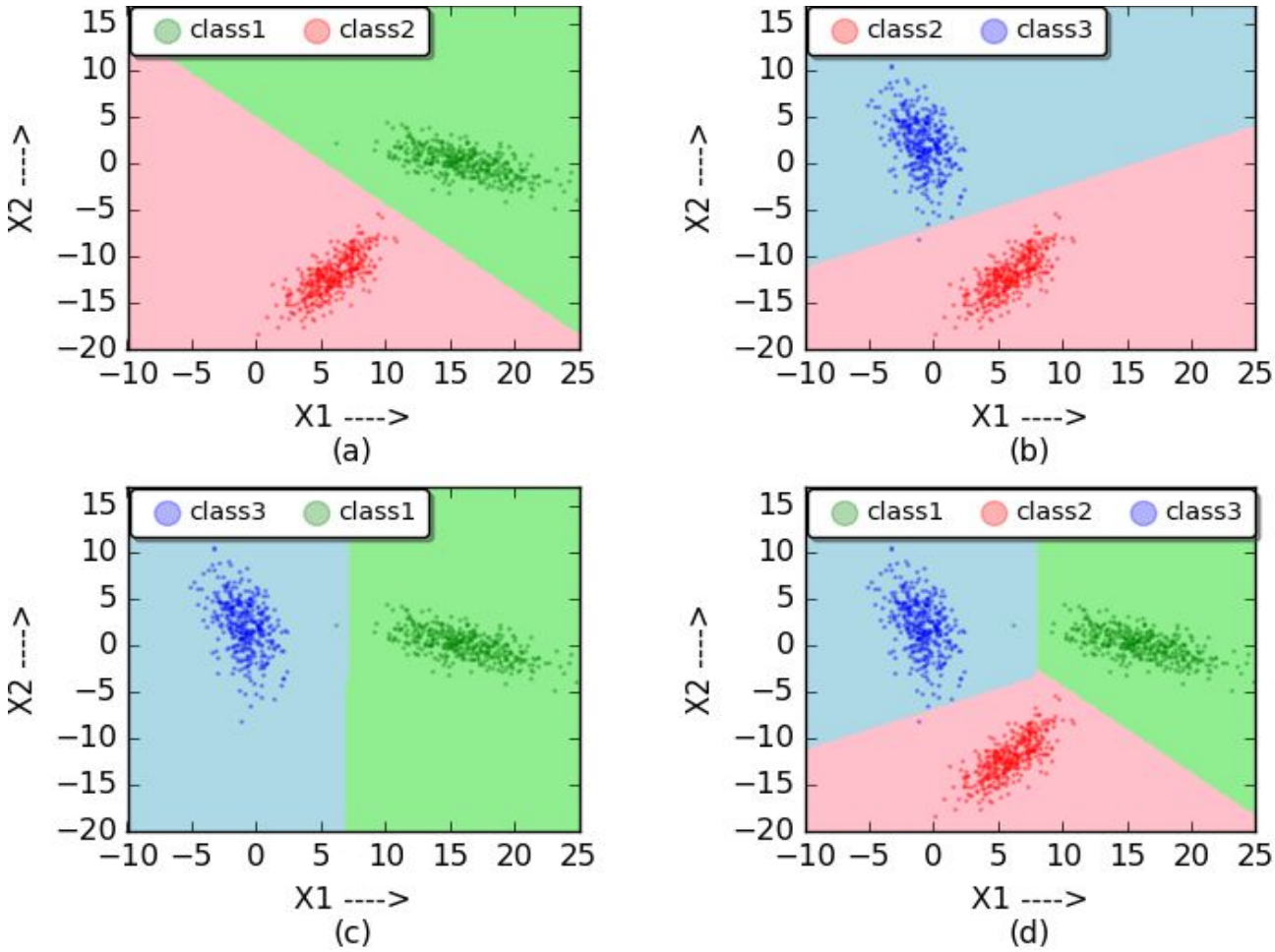

Fig 3 : Class decision boundary plot for case 2, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 99.73 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	125	0	0
Class 2	0	125	0
Class 3	0	1	124

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	1.0000	1.0000	0.9920	0.9973
Precision	1.0000	0.9921	1.0000	0.9973
F-Measure	1.0000	0.9960	0.9960	0.9973

c) Case 3 : Covariance matrix is diagonal & is different for each class

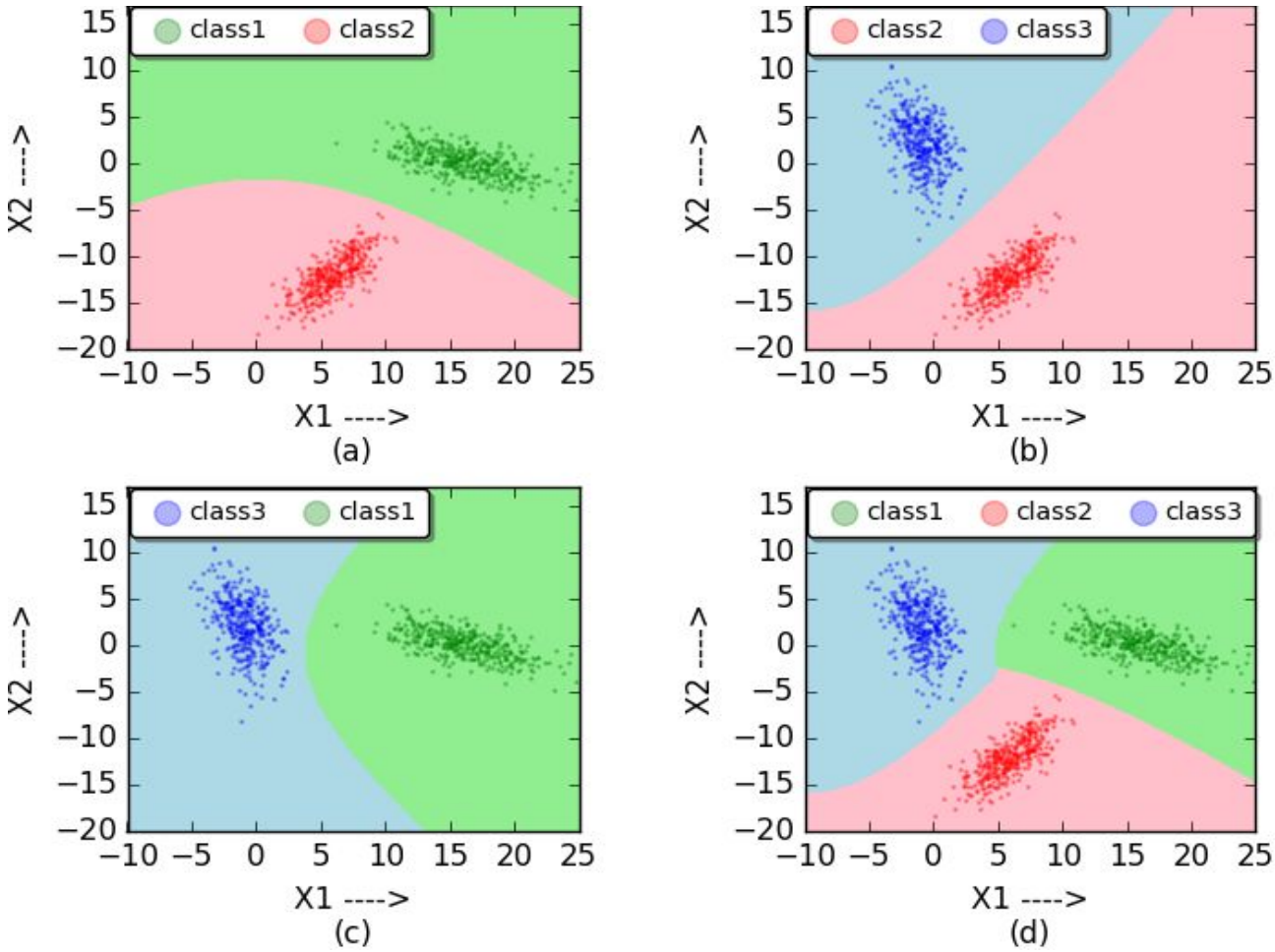

Fig 4 : Class decision boundary plot for case 3, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 99.73 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	125	0	0
Class 2	0	125	0
Class 3	0	1	124

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	1.0000	1.0000	0.9920	0.9973
Precision	1.0000	0.9921	1.0000	0.9973
F-Measure	1.0000	0.9960	0.9960	0.9973

d) Case 4 : Full Covariance matrix for each class is different.

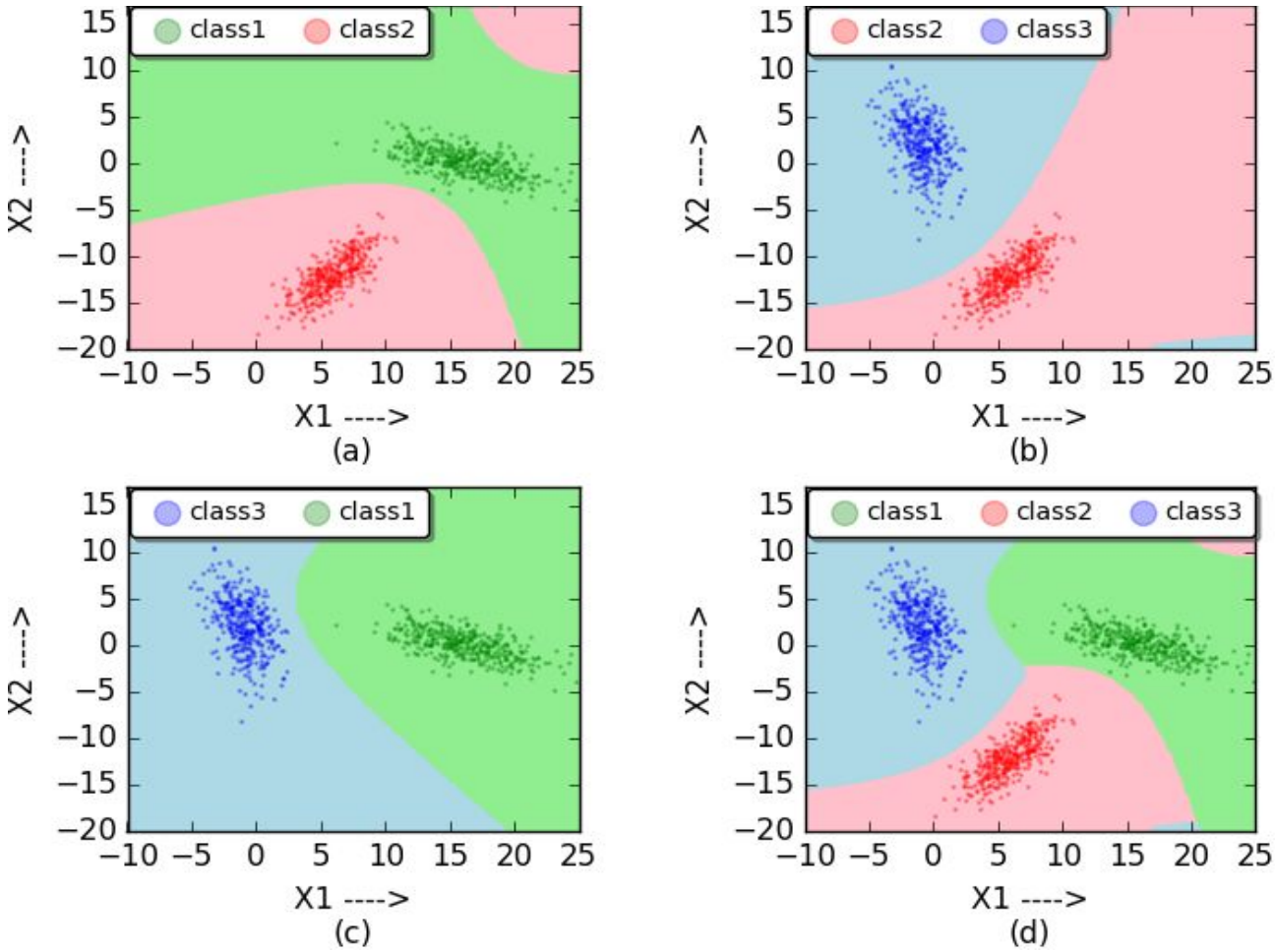

Fig 5 : Class decision boundary plot for case 4, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 100 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	125	0	0
Class 2	0	125	0
Class 3	0	0	125

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	1.0000	1.0000	1.0000	1.0000
Precision	1.0000	1.0000	1.0000	1.0000
F-Measure	1.0000	1.0000	1.0000	1.0000

3.2. Dataset II : Non-Linearly Separable Data

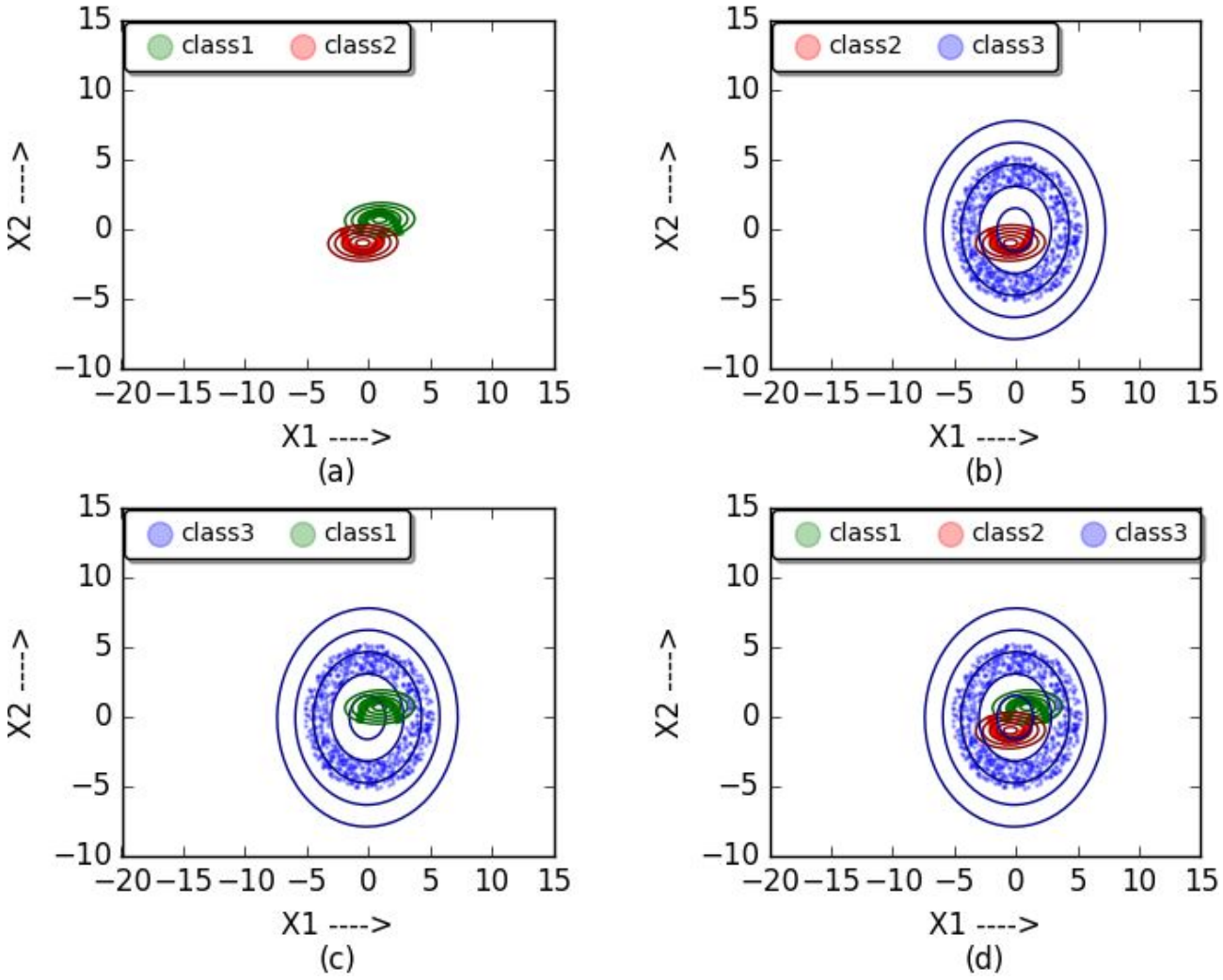

Fig 6 : Contour plot for Non-linearly separable dataset. Contour plot between : (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together

a) Case 1 : Covariance matrix for all the classes is the same and is $\sigma^2 I$

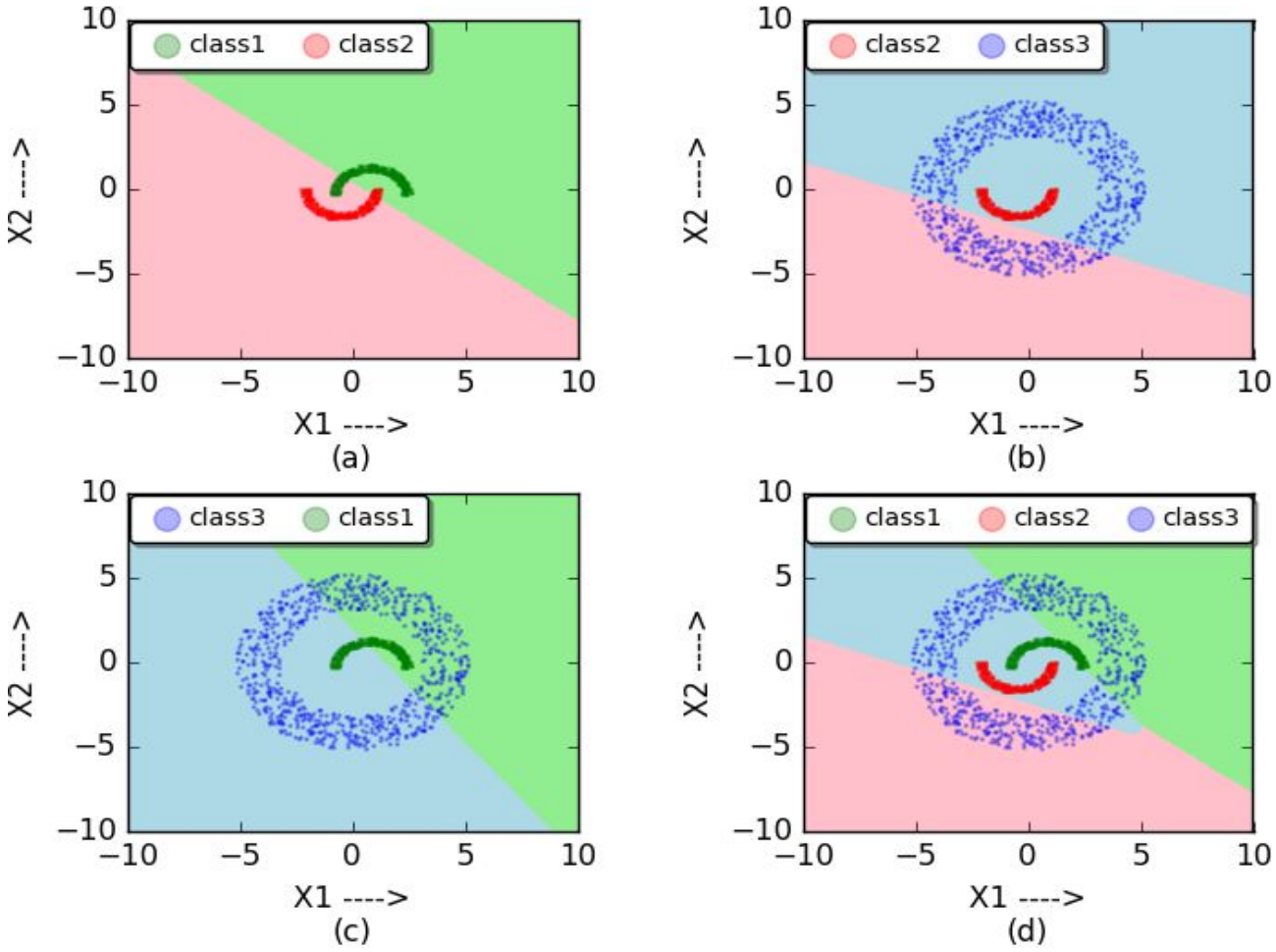

Fig 7 : Class decision boundary plot for case 1, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 26.4 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	66	0	59
Class 2	0	6	119
Class 3	109	81	60

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.5280	0.0480	0.2400	0.2720
Precision	0.3771	0.0690	0.2521	0.2327
F-Measure	0.4400	0.0566	0.2459	0.2475

b) Case 2 : Full Covariance matrix for all classes is the same and is Σ

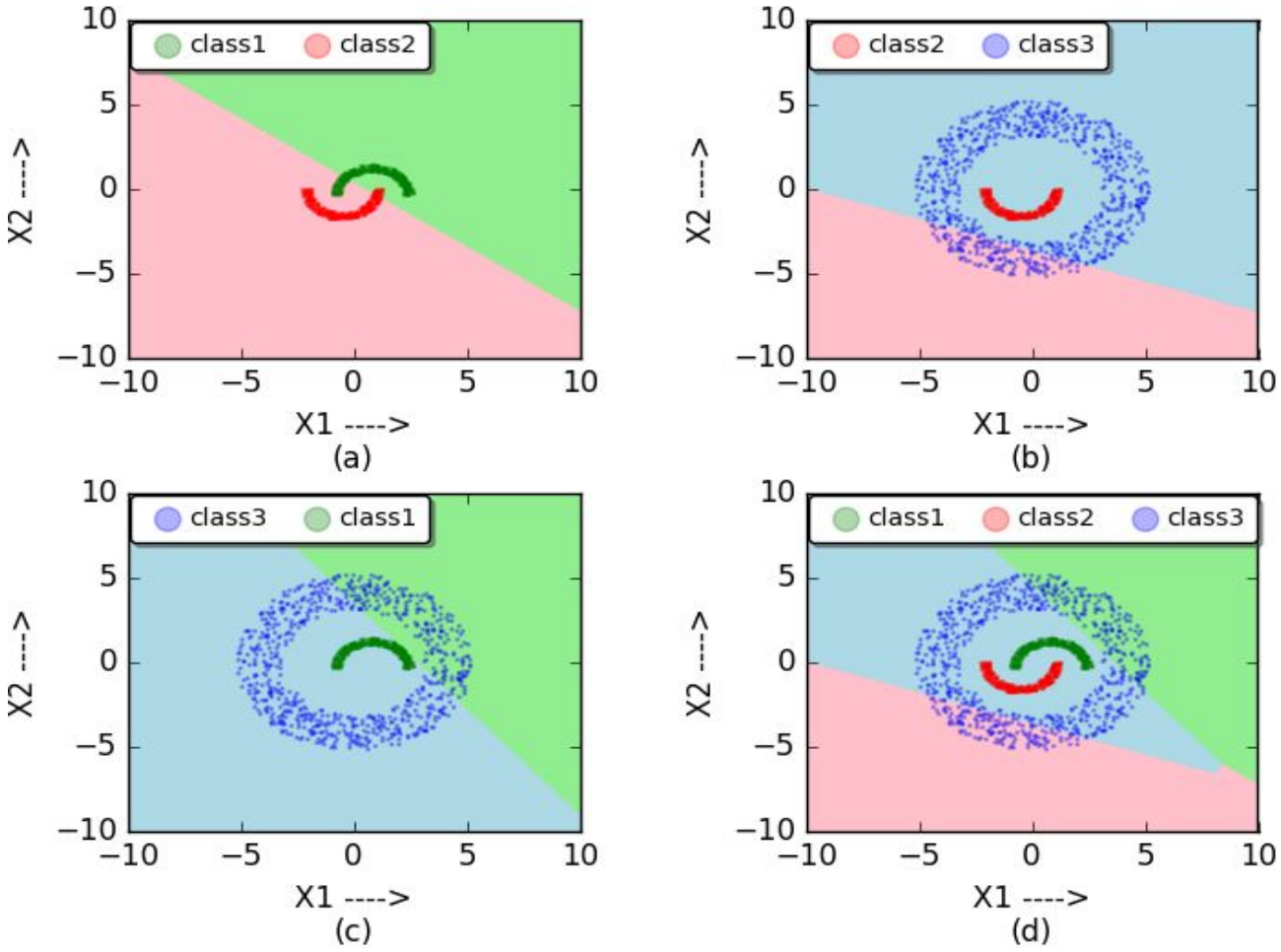

Fig 8 : Class decision boundary plot for case 2, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 22.4 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	0	0	125
Class 2	0	0	125
Class 3	89	49	112

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.0000	0.0000	0.4480	0.1493
Precision	0.0000	0.0000	0.3094	0.1031
F-Measure	0.0000	0.0000	0.3660	0.1220

c) Case 3 : Covariance matrix is diagonal & is different for each class

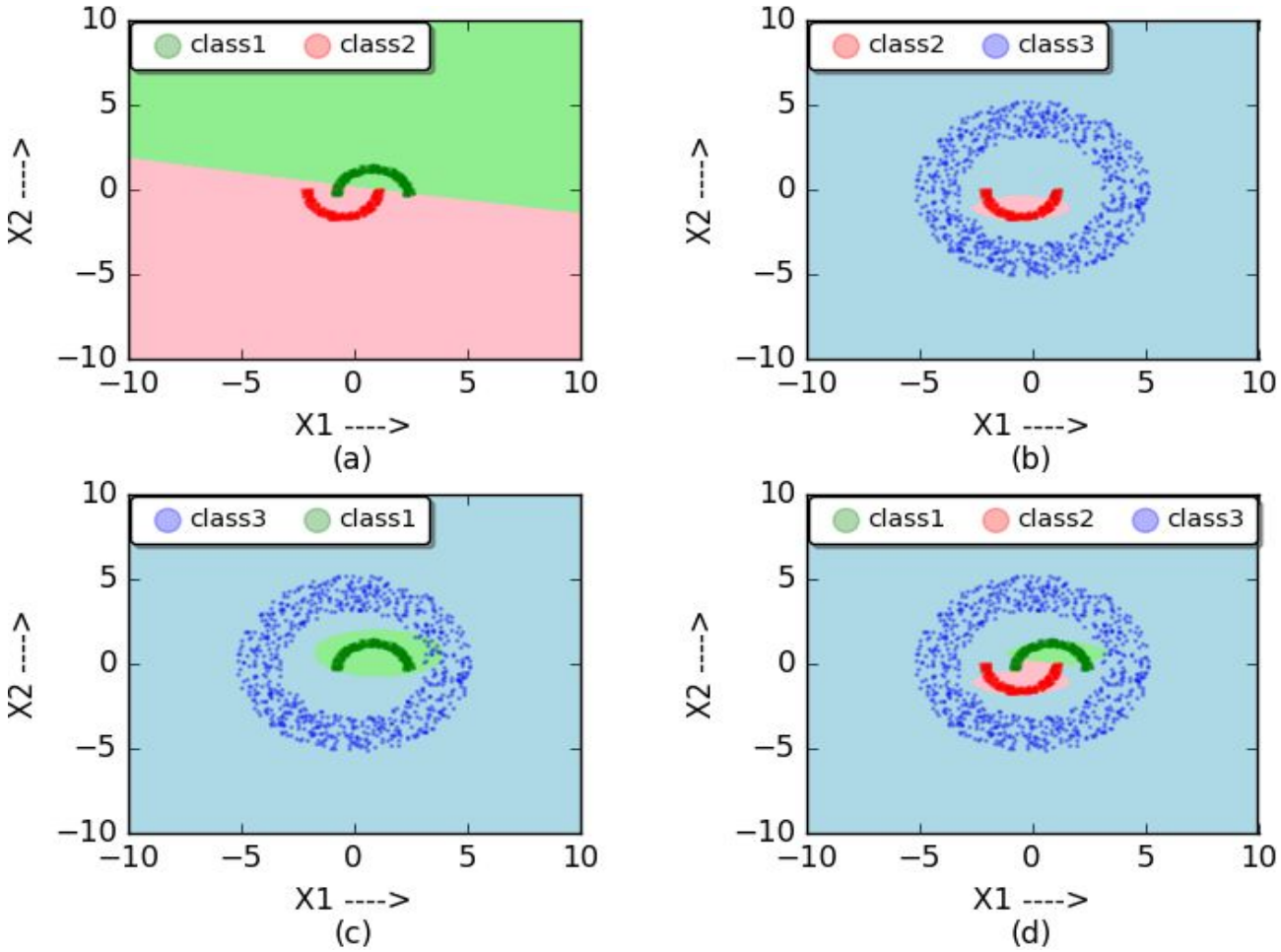

Fig 9 : Class decision boundary plot for case 3, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 94.8 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	116	4	5
Class 2	6	112	7
Class 3	4	0	246

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.9280	0.8960	0.9840	0.9360
Precision	0.9206	0.9655	0.9534	0.9465
F-Measure	0.9243	0.9294	0.9685	0.9407

d) Case 4 : Full Covariance matrix for each class is different.

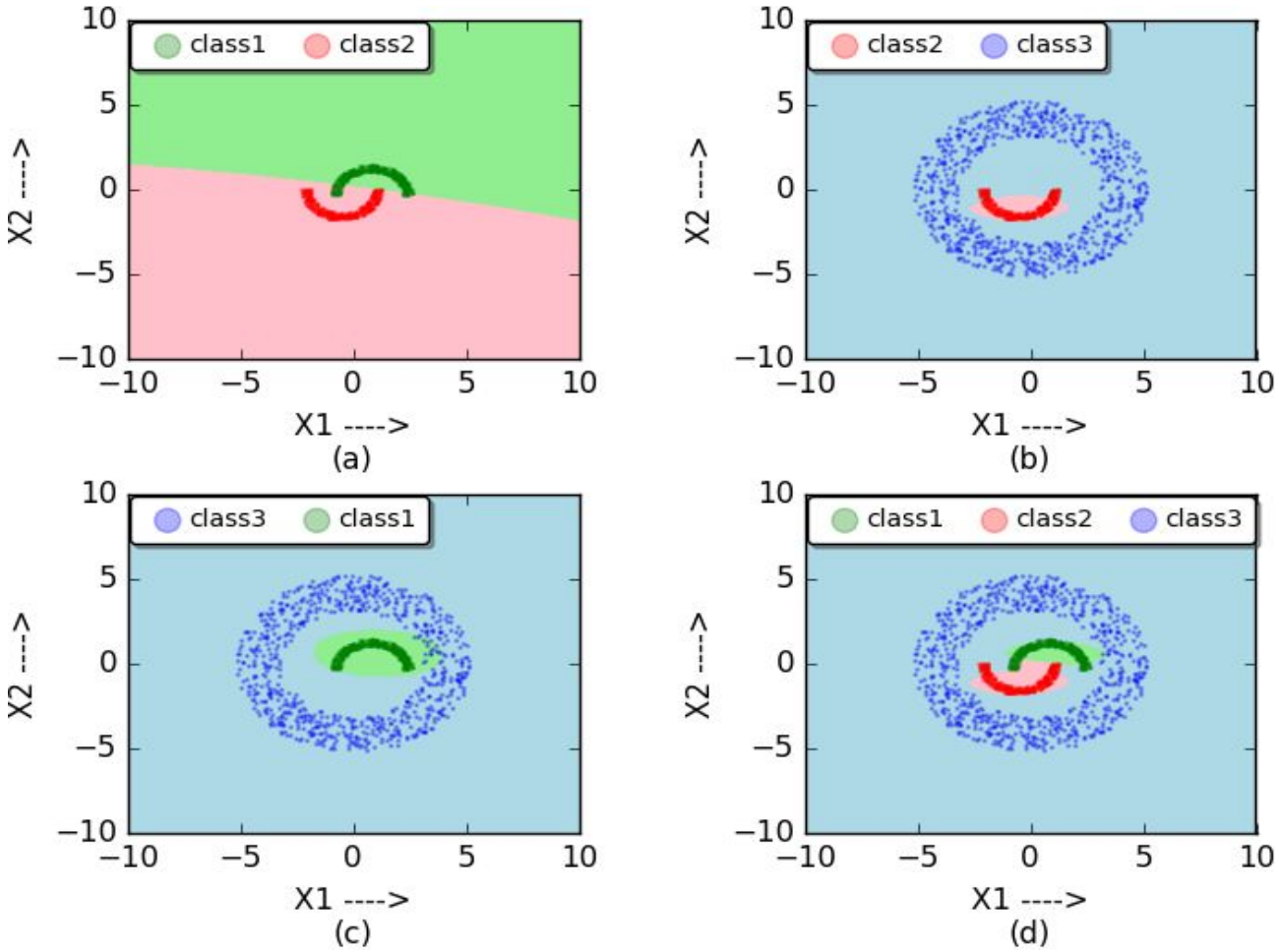

Fig 10 : Class decision boundary plot for case 4, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 94.4 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	115	5	5
Class 2	4	111	10
Class 3	4	0	246

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.920	0.8880	0.9840	0.9306
Precision	0.9349	0.9568	0.9425	0.9447
F-Measure	0.9274	0.9211	0.9628	0.9371

3.3. Dataset III : Real World Data

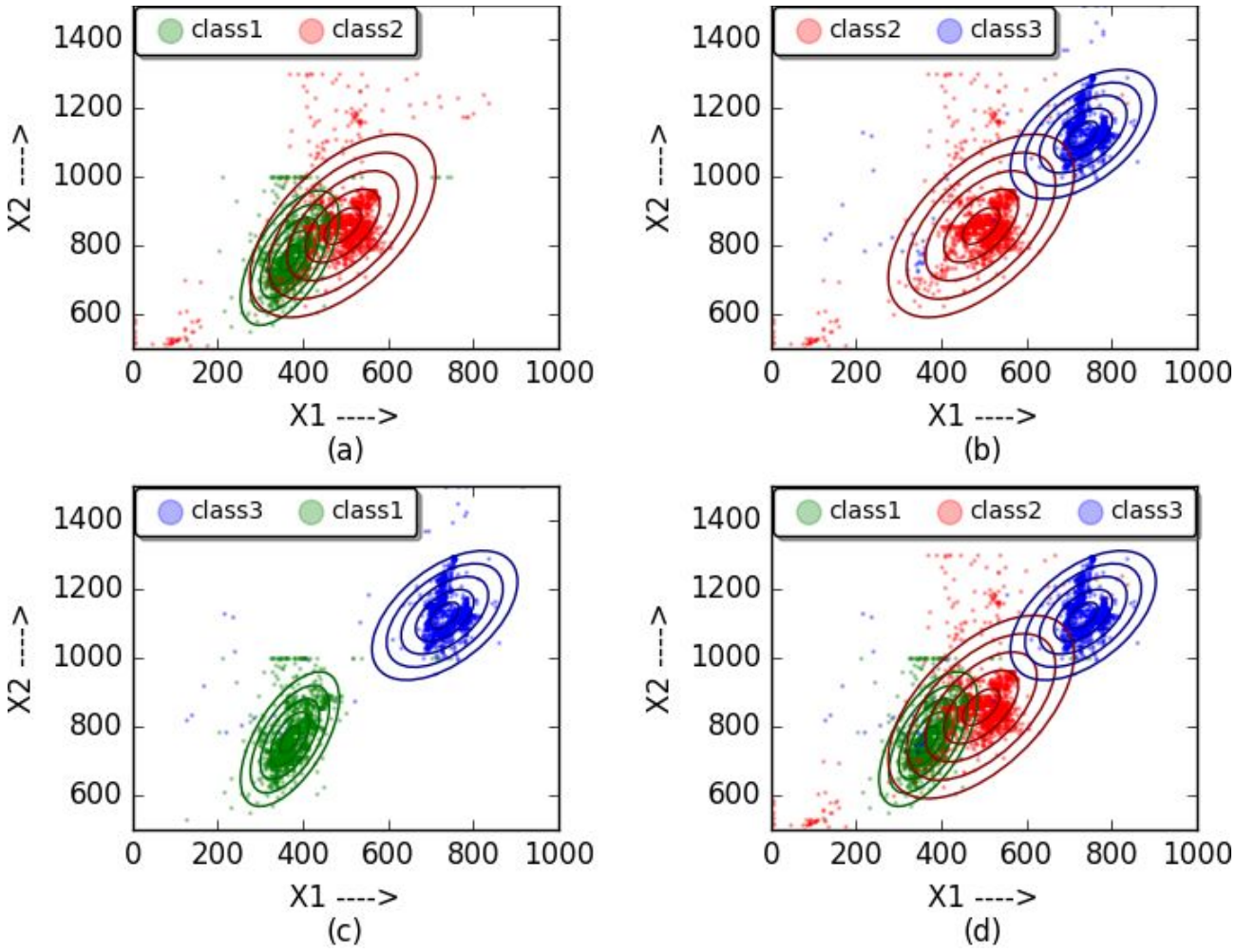

Fig 11 : Contour plot for Real World dataset. Contour plot between : (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together

a) Case 1 : Covariance matrix for all the classes is the same and is $\sigma^2 I$

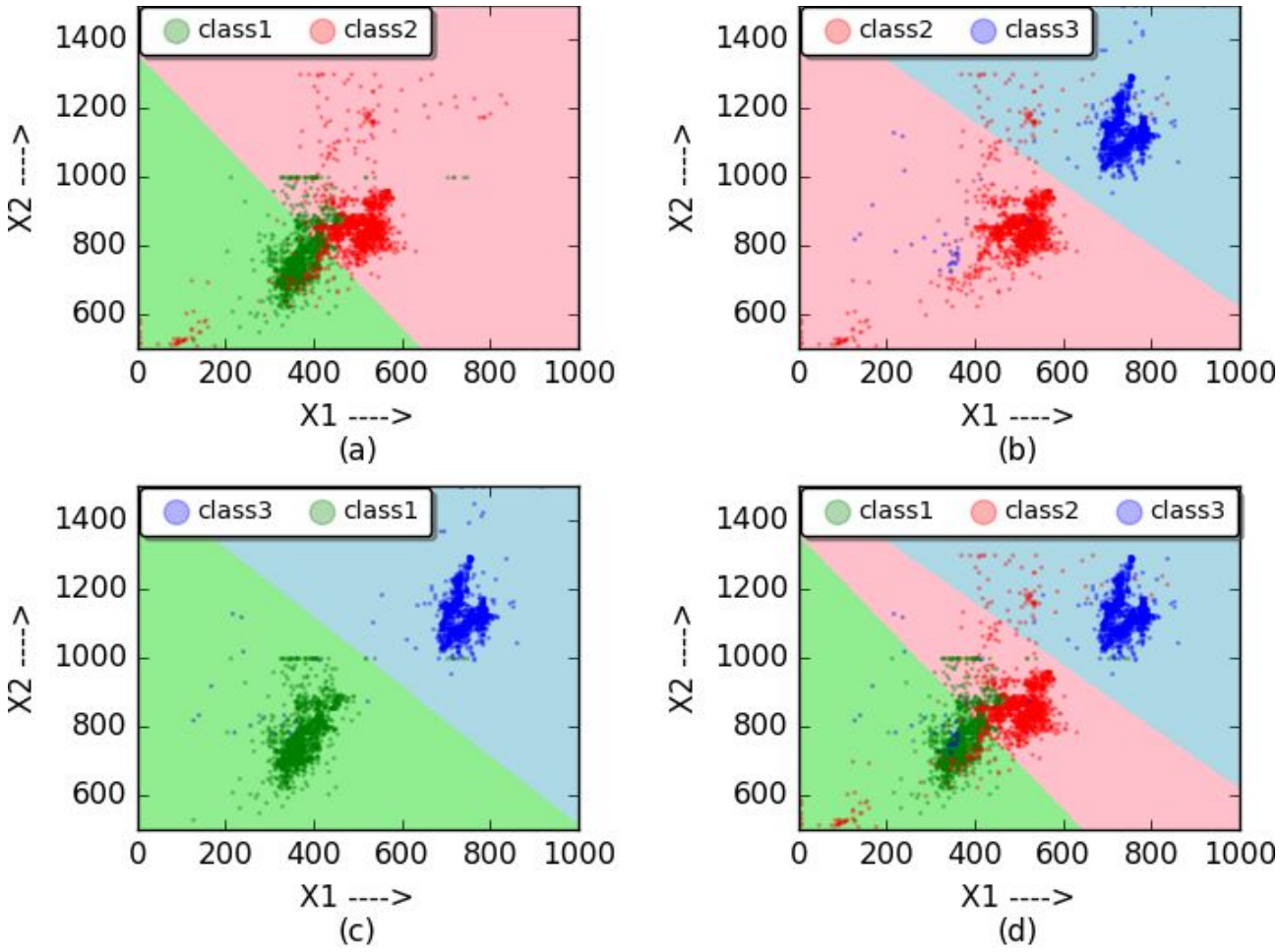

Fig 12 : Class decision boundary plot for case 1, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 83.3896 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	572	43	7
Class 2	172	397	44
Class 3	24	5	512

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.9196	0.6476	0.9463	0.8378
Precision	0.7447	0.8921	0.9094	0.8487
F-Measure	0.8230	0.7504	0.9275	0.8336

b) Case 2 : Full Covariance matrix for all classes is the same and is Σ

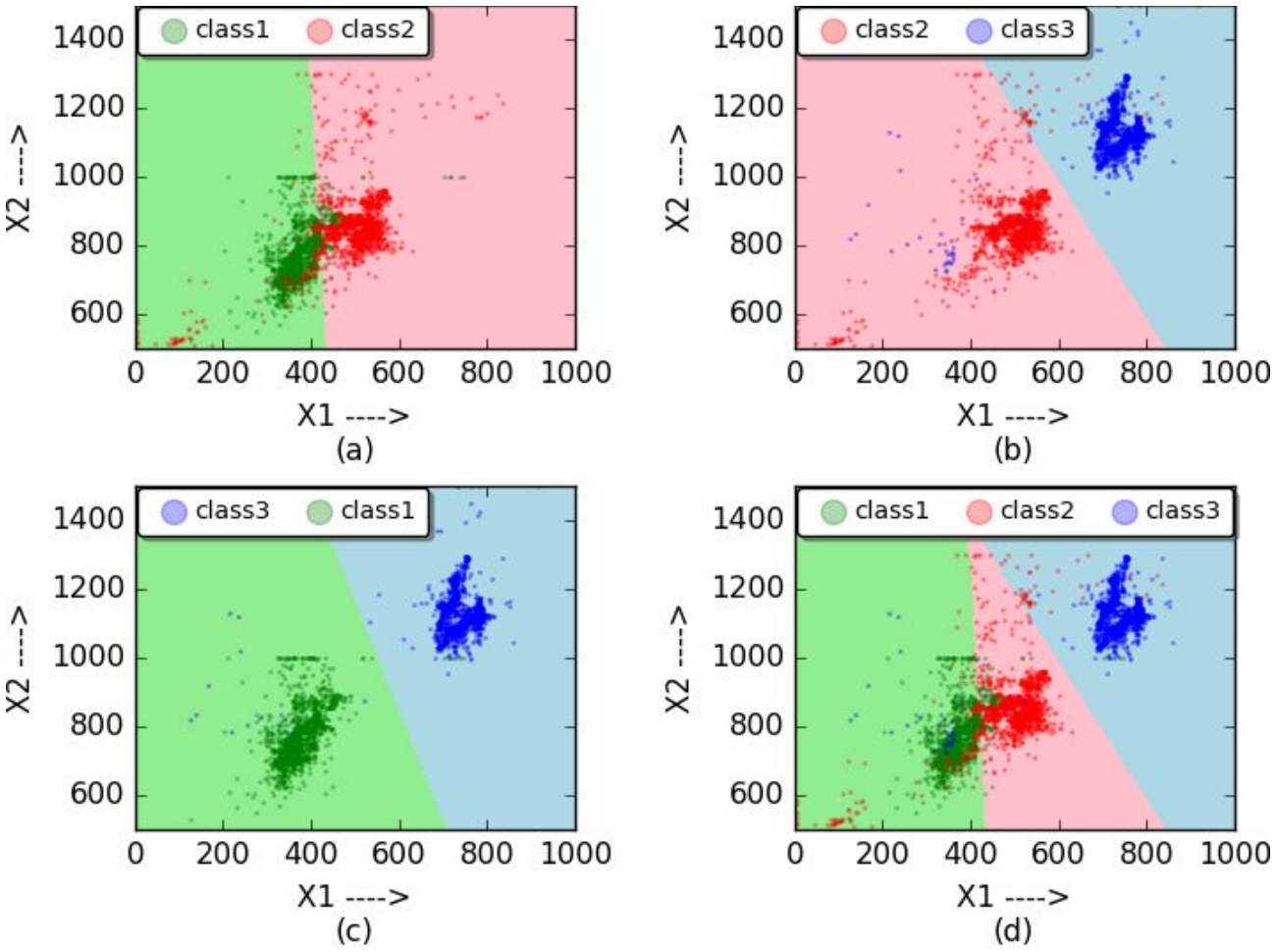

Fig 13 : Class decision boundary plot for case 2, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 84.6846 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	602	13	7
Class 2	202	392	19
Class 3	26	5	510

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.9678	0.6394	0.9426	0.8500
Precision	0.7253	0.9560	0.9514	0.8776
F-Measure	0.8292	0.7663	0.9470	0.8475

c) Case 3 : Covariance matrix is diagonal & is different for each class

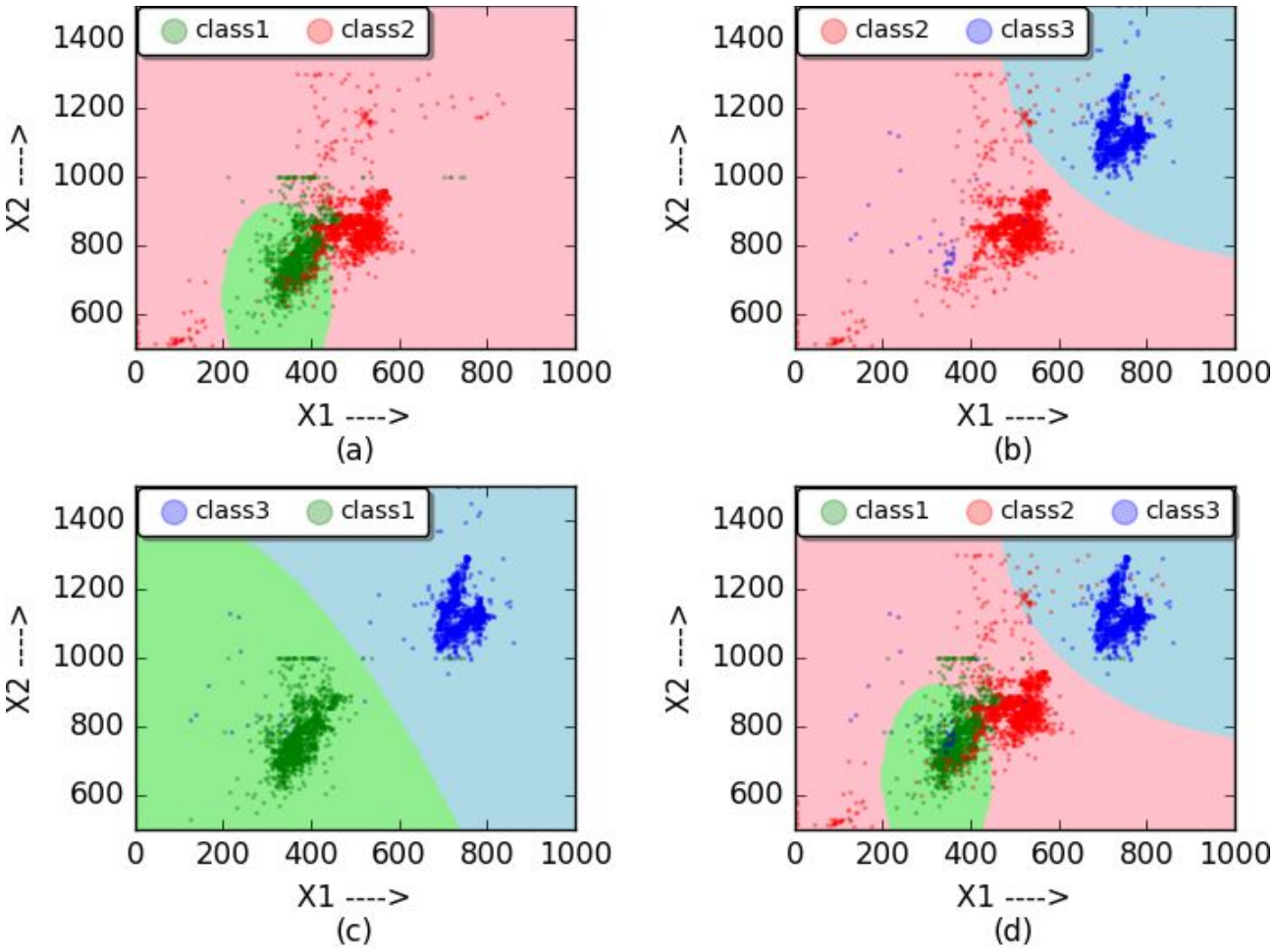

Fig 14 : Class decision boundary plot for case 3, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 82.5450 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	546	69	7
Class 2	185	410	18
Class 3	15	16	510

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.8778	0.6688	0.9426	0.8297
Precision	0.7319	0.8282	0.9532	0.8378
F-Measure	0.7982	0.7400	0.9479	0.8287

d) Case 4 : Full Covariance matrix for each class is different.

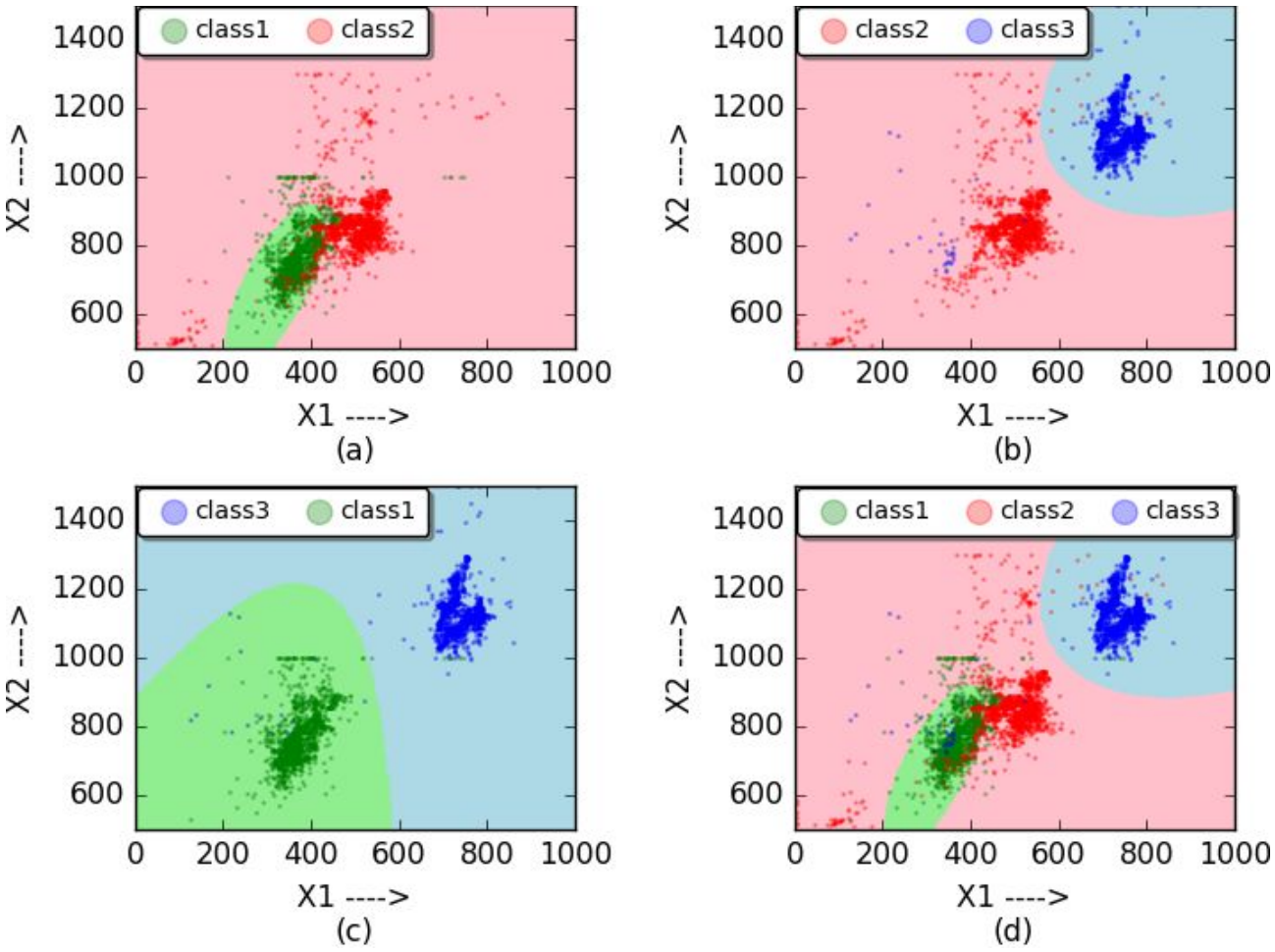

Fig 15 : Class decision boundary plot for case 4, implying on dataset 1. Decision boundary plot between (a) class1 & class2, (b) class2 & class3, (c) class3 & class1 (d) class1, class2, class3 together.

Accuracy : 81.0810 %

Confusion Matrix

	Class 1	**Class 2**	**Class 3**
Class 1	531	84	7
Class 2	208	400	5
Class 3	13	19	509

Analysis

	Class 1	**Class 2**	**Class 3**	**Mean**
Recall	0.8536	0.6525	0.9408	0.8156
Precision	0.7061	0.7952	0.9769	0.8261
F-Measure	0.7729	0.7168	0.9585	0.8161

4. Conclusion

1) For linearly separable artificial data, Bayes classifier works excellent with very high accuracy (close to 100%) for all cases.
2) For non-linearly separable artificial data, Bayes classifier fails for case 1 and case 2 but works well for case 3 and case 4.
3) For real-world data results were good for all cases but not as good as given artificial data because of overlapping data.
4) For case 1(covariance matrix = $\sigma^2 I$) & case 2(covariance matrix = Σ), the decision boundary is linear for all datasets because the discriminant function will be a linear equation.
5) For case 3 and case 4, the decision boundary is non-linear for all datasets because the discriminant function will be a hyper quadratic equation.
6) As it can be seen, performance of classifier changes with change in choice of covariance matrix. For the same dataset, accuracy is different for different choice of covariance matrix.