EE 565 Machine Learning Project 2 Report

Mehrdad Ghyabi

***Abstract*—This is a summarized report about implementation and evaluation of three different types of classification algorithms. On top of that, these three algorithms are compared to each other. First, Least Square Classification is implemented on three different datasets consisting two and three classes of data. After that, Maximum Mean Projection Classification is implemented on the same datasets with two classes of data. Last but not least, Fisher's Discriminant algorithm is used to classify those datasets.**

***Index Terms*— Least Square Classification, Maximum Mean Projection Classification, Fisher's Discriminant.**

I. Introduction

This report is consisted of six sections. After the introduction and in section II, different datasets consisting of two classes of data are classified using Least Square Classification algorithm. In section III the same algorithm is used to classify a dataset with three classes of data. Maximum Mean Projection Classification algorithm is used in section IV to classify the same datasets from section II. In section V, the Fisher's Discriminant algorithm's ability to classify the same datasets is evaluated. The las section is a summarized conclusion about comparison of results from different classification algorithms.

II. Problem 1: Least Squares Classification (Two Classes)

A. Classification of "dataset A"

In this part, a given dataset (dataset A) containing two clusters of data is classified using Least Square Classification algorithm. A one-hot encoded target is used in order to make algorithm able to classify more than two classes of data. The one-hot encoding was implemented using Equation (1).

$$C_k = \arg max_{1\le j\le K}\, \boldsymbol{y}_j(\boldsymbol{x}) \tag{1}$$

In which $\boldsymbol{y}$ is discriminant function in Equation (2).

$$\boldsymbol{y} = \boldsymbol{XW} \tag{2}$$

$\boldsymbol{X}$ is the data matrix and $\boldsymbol{W}$ is the result of Equation (3).

$$\boldsymbol{W} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T T \tag{3}$$

The result of classification is presented in Figure 1. This algorithm is working perfectly classifying two classes present in dataset A.

B. Classification of "dataset A2"

The same algorithm was used to classify the second dataset and the result is presented in Figure 2. As it was anticipated, the presence of a few datapoints away from the center of one of the clusters caused the algorithm not to work as well as the last part.

C. Classification of a double moon dataset

In this part a double moon data set with 5000 data points is fed to the classifier and the result is presented in Figure 3. Parameters of the double moon dataset are presented in Table 1. Since it is impossible to separate this dataset with a straight line, some data points are classified to the wrong class.

TABLE I
PARAMETERS OF THE DOUBLE MOON DATASET

Parameter	Value
N	5000
d	-0.1
r	1
w	0.6

III. Problem 2: Least Squares Classification (Three Classes)

A given dataset (Dataset B) was fed to the classifier model from last section and the result is presented in Figure 4. The algorithm worked perfectly for this dataset, dividing the domain into three convex decision regions, meaning that a line connecting two arbitrary points in a give region remain in the same region.

IV. Problem 3: Maximum Mean Projection Classification

A. Classification of "dataset A"

In this part, dataset A containing two clusters of data is classified using Maximum Mean Projection Classification algorithm. To implement this algorithm, first mean of each class is calculates ($\boldsymbol{m}_1$ and $\boldsymbol{m}_2$). Then $\boldsymbol{w}$ is resulted from Equation (4).

$$\boldsymbol{w} = \frac{\boldsymbol{m}_2 - \boldsymbol{m}_1}{\|\boldsymbol{m}_2 - \boldsymbol{m}_1\|} \tag{4}$$

Then the data is classified using the threshold y_0 from Equation (5).

$$\boldsymbol{y_0} = \boldsymbol{mean}(\boldsymbol{Xw}) - w_1 \tag{5}$$

The resulting classification is presented in Figure 5. The data points are illustrated with their original class marks in order to show the error of this classifier. The histogram of resulting classification is shown in Figure 6.

B. *Classification of "dataset A2"*

The same algorithm was used to classify the second dataset and the result is presented in Figure 7 and the histogram of classified datapoints in Figure 8. As it was expected, the presence of a few outlier data points did not affect the precision of the classifier by much.

C. *Classification of a double moon dataset*

A double moon data set with the same parameters as in section II was fed to this classifier and results and their histograms are presented in Figure 9 and Figure 10 respectively. In case of the double moon dataset, the classifier in section II works better. In other words, Maximum Mean Projection classification is not helping with the nonlinearity of the decision boundary.

V. Problem 4: Fisher's Discriminant

A. *Classification of "dataset A"*

In this part, dataset A containing two clusters of data is classified using Fisher's Discriminant algorithm. The procedure is quite analogous to that of section IV. The only difference is that here is different.

$$\boldsymbol{w} = \frac{\boldsymbol{S}_w^{-1}(\boldsymbol{m}_2 - \boldsymbol{m}_1)}{\|\boldsymbol{S}_w^{-1}(\boldsymbol{m}_2 - \boldsymbol{m}_1)\|} \tag{6}$$

In which $\boldsymbol{S}_w$ is:

$$\boldsymbol{S}_w = \sum_{n \in c_1} (\boldsymbol{x}_n - \boldsymbol{m}_1)(\boldsymbol{x}_n - \boldsymbol{m}_1)^T + \sum_{n \in c_2} (\boldsymbol{x}_n - \boldsymbol{m}_2)(\boldsymbol{x}_n - \boldsymbol{m}_2)^T \tag{7}$$

The classified domain and histograms of classified datapoints are shown in Figure 11 and Figure 12 respectively.

B. *Classification of "dataset A2"*

The same algorithm was used to classify the second dataset and the result is presented in Figure 13 and the histogram of classified datapoints in Figure 14.

C. *Classification of a double moon dataset*

A double moon data set with the same parameters as in section II was fed to this classifier and results and their histograms are presented in Figure 15 and Figure 16 respectively. In general performance of algorithms in Sections IV and V do not seem to be much different.

VI. Conclusions

In case of datasets in which different classes are well-clustered least square classification has a better performance. In datasets in which classes contain outlier datapoints maximum mean projection and fisher's discriminant seem to be more promising.

When nonlinear boundaries are needed to separate classes, like the case of double moon dataset, least square classification shows better results.

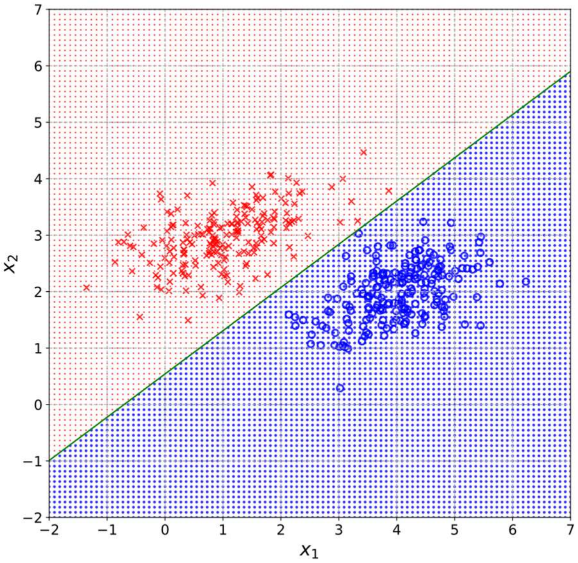

Fig. 1 Result of classifying "dataset A" by least square classification algorithm

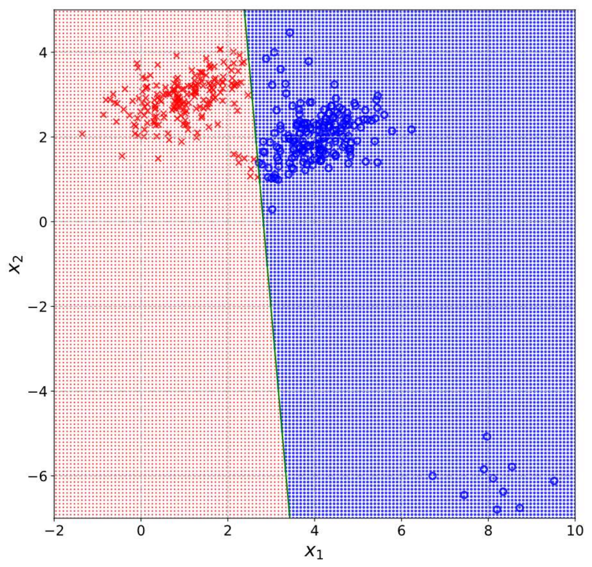

Fig. 2 Result of classifying “dataset A2” by least square classification algorithm

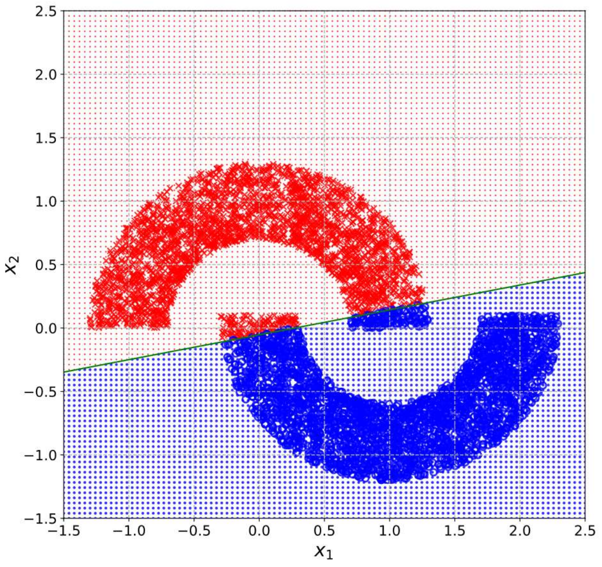

Fig. 3 Result of classifying a double moon dataset by least square classification algorithm

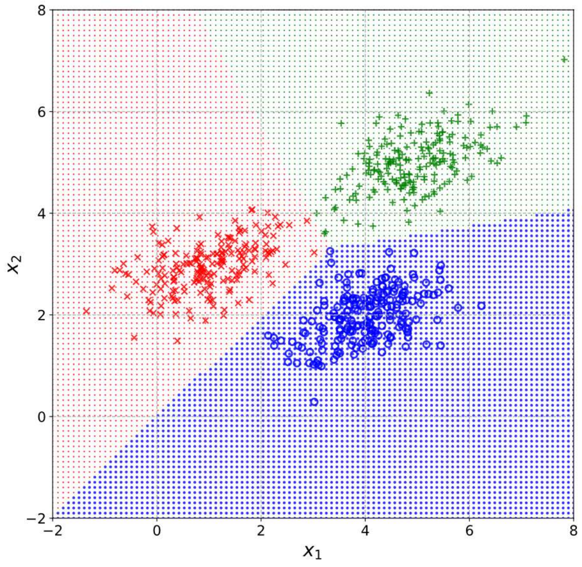

Fig. 4 Result of classifying “dataset B” by least square classification algorithm

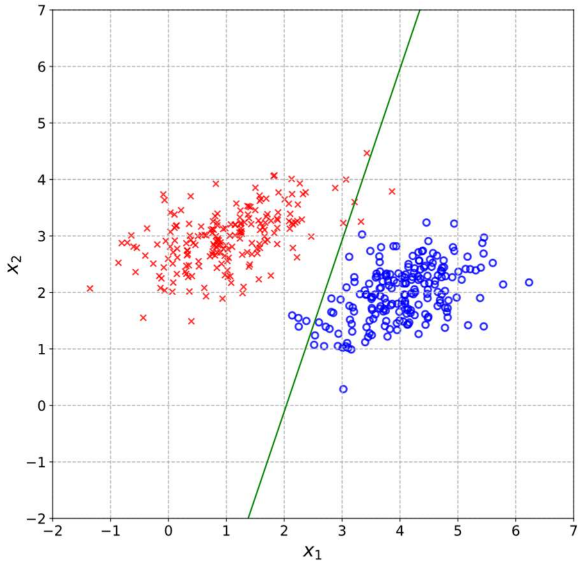

Fig. 5 Result of classifying “dataset A” by maximum mean projection classification algorithm

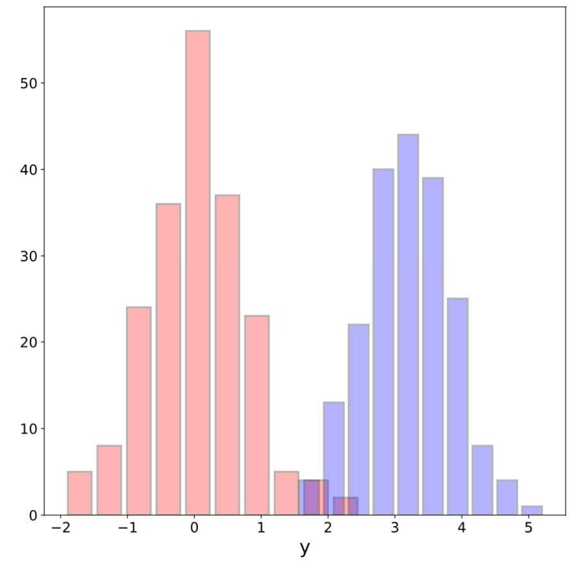

Fig. 6 Histogram of classified dataset A using maximum mean projection classification algorithm

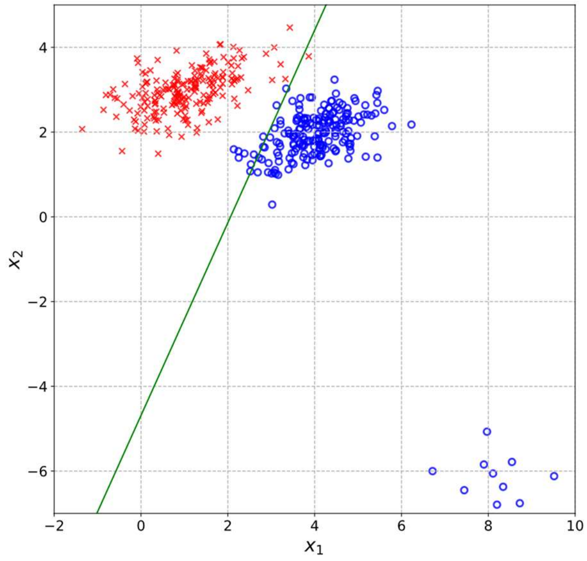

Fig. 7 Result of classifying "dataset A2" by maximum mean projection classification algorithm

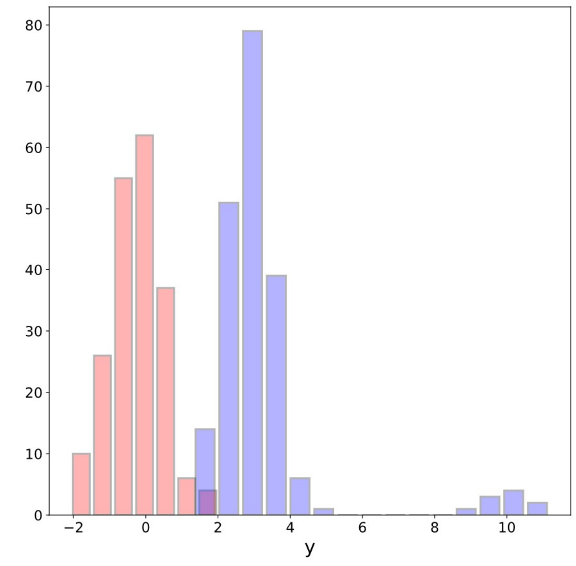

Fig. 8 Histogram of classified dataset A2 using maximum mean projection classification algorithm

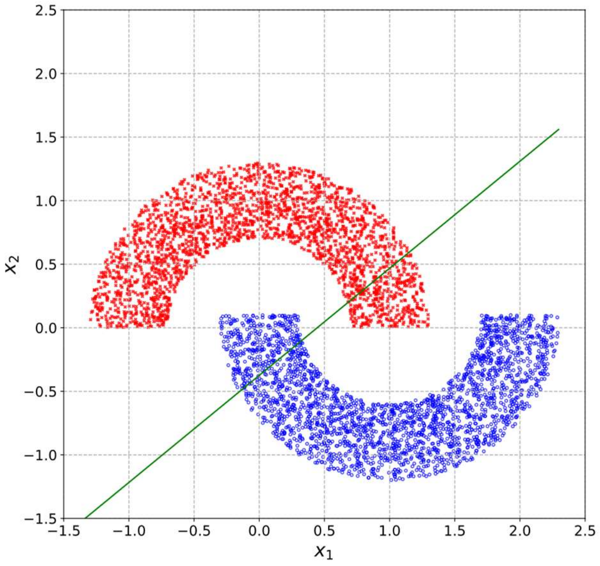

Fig. 9 Result of classifying double moon dataset by maximum mean projection classification algorithm

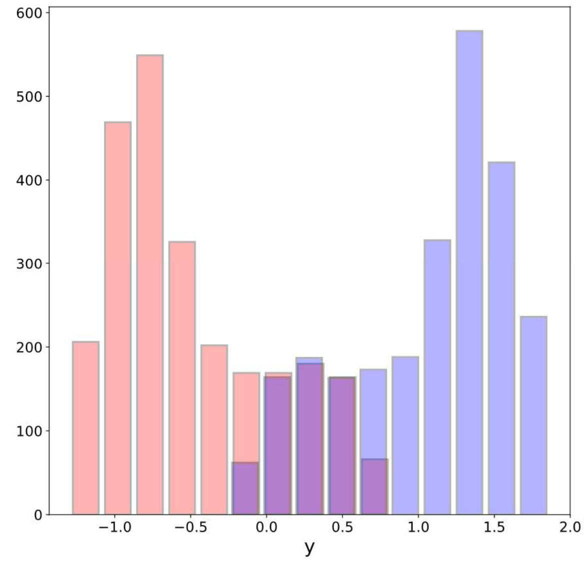

Fig. 10 Histogram of classified double moon dataset using maximum mean projection classification algorithm

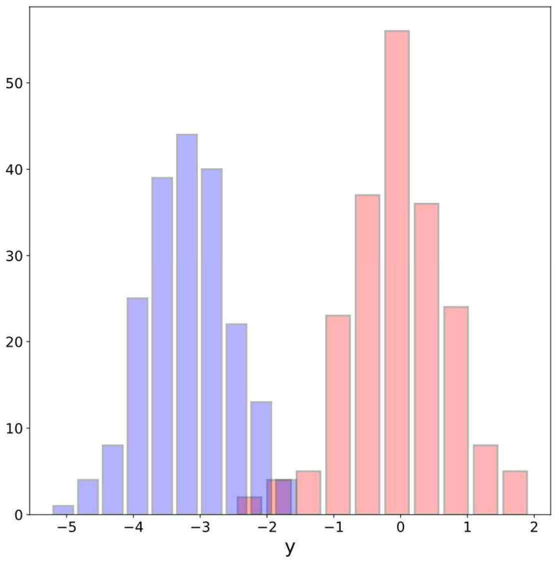

Fig. 12 6 Histogram of classified dataset A using fisher's discriminant algorithm

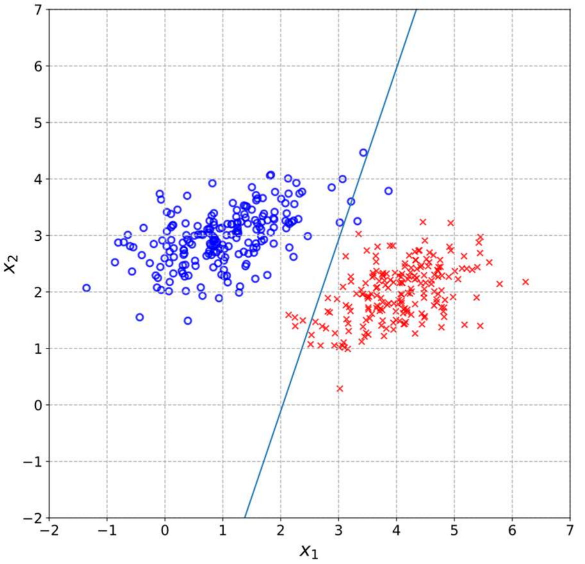

Fig. 11 Result of classifying "dataset A" by fisher's discriminant algorithm

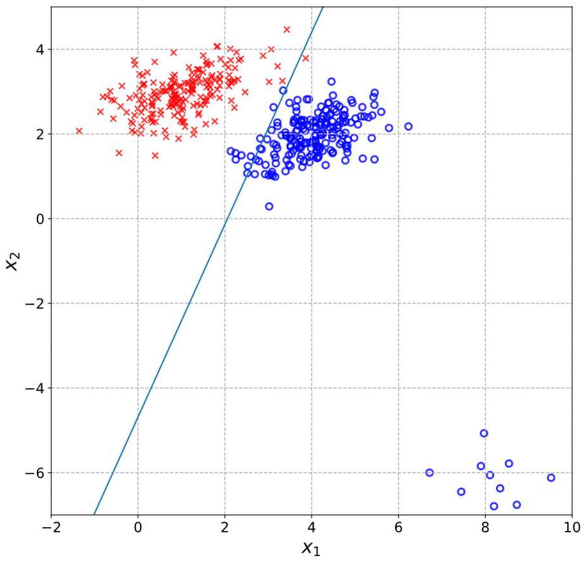

Fig. 13 Result of classifying "dataset A2" by fisher's discriminant algorithm

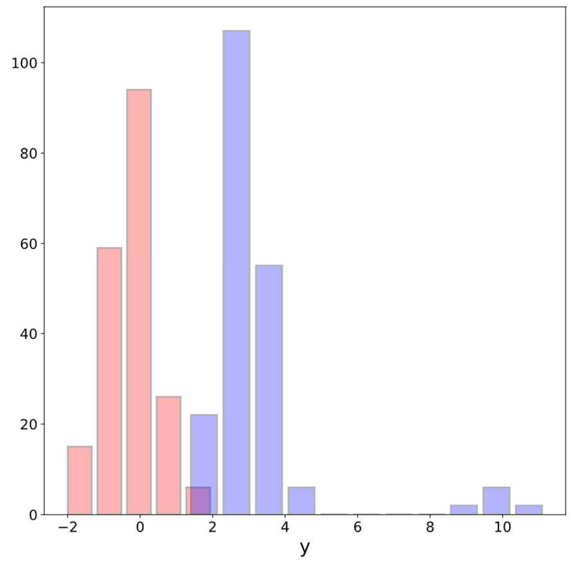

Fig. 14 Histogram of classified dataset A2 using fisher's discriminant algorithm

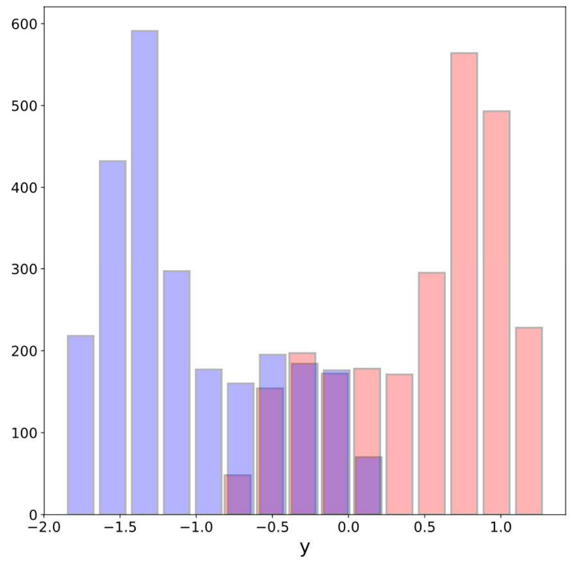

Fig. 16 Histogram of classified double moon dataset using fisher's discriminant algorithm

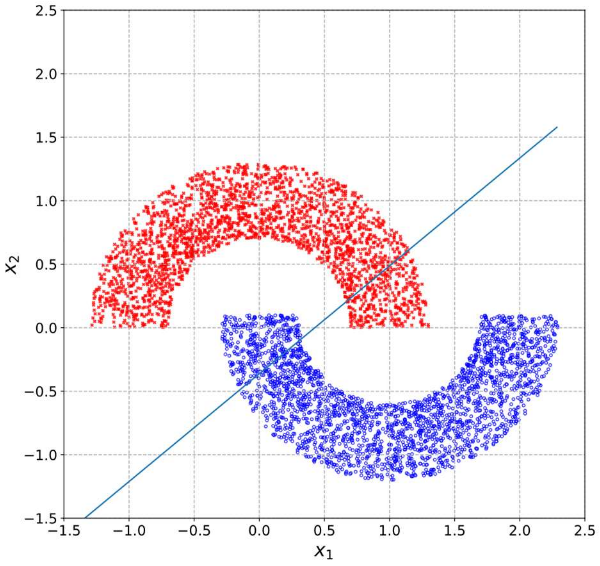

Fig. 15 Result of classifying double moon dataset by fisher's discriminant algorithm