Machine Learning 2023 spring

HW2: Classification

Deadline: 112.4.19

Part I.

Consider that there are a group of college basketball players, who come from four different high schools. Now we analyze the Performance-Rating of "skill" and the "athleticism" for every player, as shown in Figure 1, where the blue, green, red, and black dots represent the players from High school 1, High school 2, High school 3, and High school 4, respectively. Noted that there are 400, 250, 200, and 150 players from High school 1, High school 2, High school 3, and High school 4, respectively.

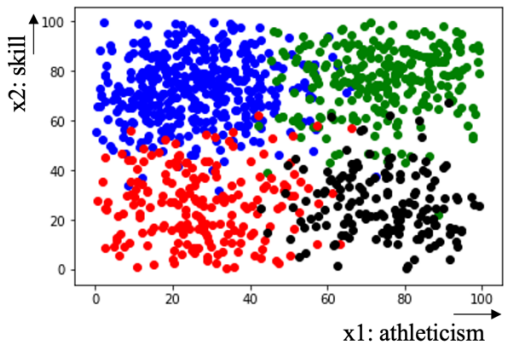

Figure 1

You are given a HW2.xlsx file. Please implement the algorithms of the generative model and the discriminative model to classify the data and plot the corresponding decision boundaries, like Figure 2.

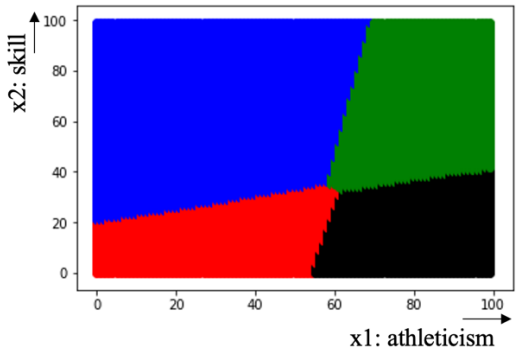

Figure 2

Key:

For the generative model, you can assume the data comes from four Gaussian distribution with the same covariance matrix but different mean vectors, and then try to find the parameters of these Gaussian distribution according to the given data. On the contrary, for the discriminative model, you do not have any assumption of the distribution for the given data. Instead, you define lots of basis functions and just try

to learn the corresponding weights for every basis function according to the given data. Since the close-form solution of the optimal weights does not exist, you need to sequentially adapt the weights based on the gradient decent method. Specifically, you should implement the Iterative reweighted least squares algorithm (IRLS) to learn the optimal weights. Noted that, for the generative model, you only need to plot one decision boundary. On the other hand, for discriminative model, you need to plot several decision boundaries for different weights in the learning process.

Part II.

According to the overall Performance-Rating of "skill" and the "athleticism" for each high school in Figure 1, we classify High school 2 to be class A, High school 1 and High school 4 to be class B, and High school 3 to be class C, as illustrated in green, blue, and red dots, respectively in Figure 3.

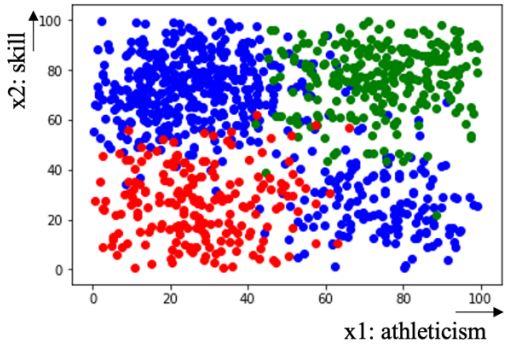

Figure 3

Please classify the data in Figure 3 for the same requirement in **Part I**.

Homework Rules and Grading Policy

Homework will be graded by:

1. The correctness of your classification result.
2. Your discussion of the comparison of the generative model and the discriminative model.

Upload:

[web] E3

[File Name] hw2_StudentID.zip (ex: hw2_1234567.zip)

Remind:

1. Your report in the format of .pdf.
2. Deadline:

   If you have a late submission by 1 to 7 days, you will only get 70% of the score. We DO NOT accept any late submission after 7 days after the deadline.