PGP in AI/ML

Feature Engineering -Assignment 3

Total Marks: 17

The following questions are to enhance your understanding of basic concepts and definitions. You are expected to answer these questions on your own without referring to internet.

Q1. Aggregation and PCA perform reduce dimensions. What are the differences between both techniques? [2 M]

Note: Aggregation is groups of dimensions are combined. In some cases, as when days are aggregated into months or the sales of a product are aggregated by store location, the aggregation can be viewed as a change of scale.
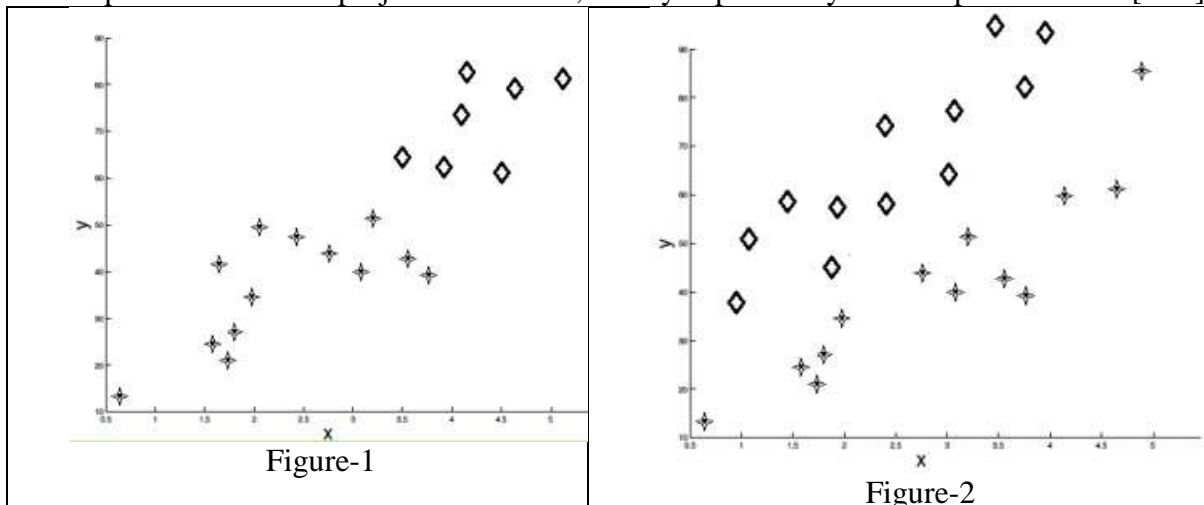
Q2. What does a smaller value of eigenvalue indicate in principal component analysis? [2 M]

Q3. What are the differences between Feature Subset Selection and Dimension Reduction using PCA? [4 M]

Q4. Figure-1 and 2 shows two different data sets in R^2 with coordinates (x,y).Answer the following two questions.
   i) Draw the first two principal components. [4 M]
   ii) Once the data set is projected can we separate the two classes? If yes, which Principal Component should we project onto? If no, briefly explain why it is not possible. [5 M]

Figure-1

Figure-2

Submission Details

1. Final dataset - id_final_dataset.csv

Contact Details
You should put up queries in the discussion forum of the corresponding assignment folder only.
1. Himanshu - himanshuchawla437@gmail.com
2. Mohit - mohitrathore15@gmail.com
3. Kartheek - f20160015@hyderabad.bits-pilani.ac.in
4. Shristy - f20160688@hyderabad.bits-pilani.ac.in
a) Agam - f20160484@hyderabad.bits-pilani.ac.in

*************************** *That's all folks* ******************************