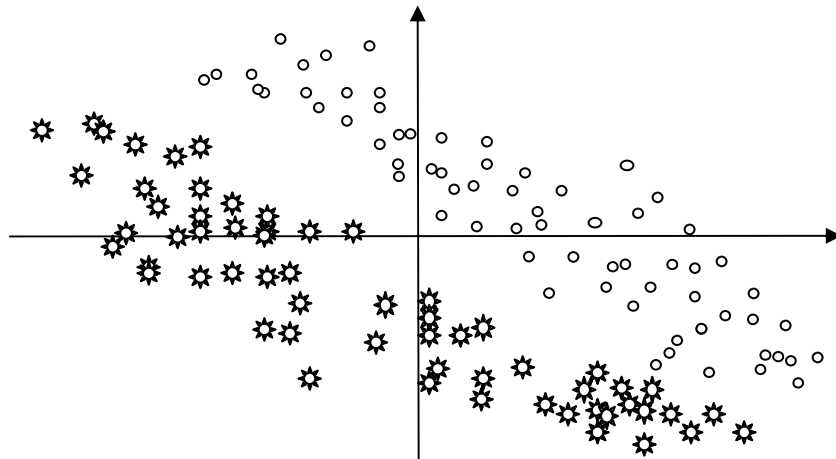


Name:

Student ID#:

Statistical Pattern Recognition (CE-725)
Department of Computer Engineering
Quiz #2 (Feature Selection & Extraction) - Spring 2012

1. (5 points) Explain problem of over-fitting in pattern classification.
2. (5 points) Explain the role of feature dimensionality in problem of over fitting?
- 3) (10 points) Suppose, for simplicity, that your input feature vectors are 2D and distributed according to the following plot. You apply PCA and only keep one eigenvector. On the figure, draw the direction of the leading eigenvector.



Next, explain briefly why, for this 2D dataset, the use of PCA for dimensionality reduction, down to 1D, actually makes classification much harder than it should be, if not impossible.