**University of Waterloo**

**Faculty of Engineering**

**SYDE 372 Lab 1 Report**

**Clusters and Classification Boundaries**

**Prepared By:**

**Oscar Lo, ID XXXXXXXX**

**Dae Jeon, ID XXXXXXXX**

**Charlotte Emily Bond, ID XXXXXXXX**

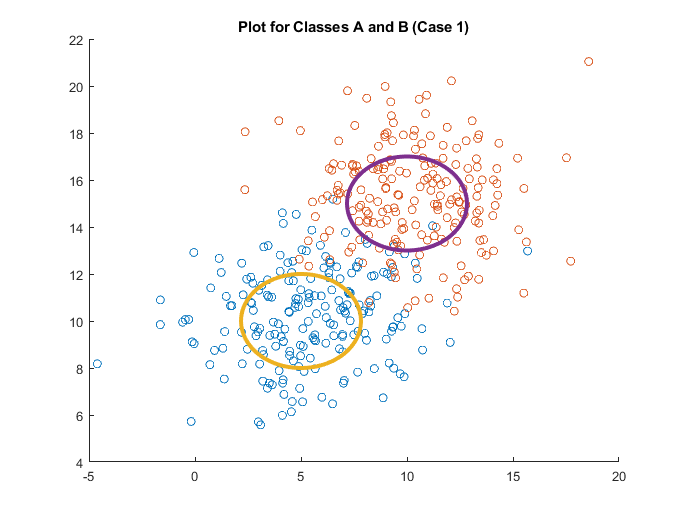
**Zixuan Ren, ID 20566221**

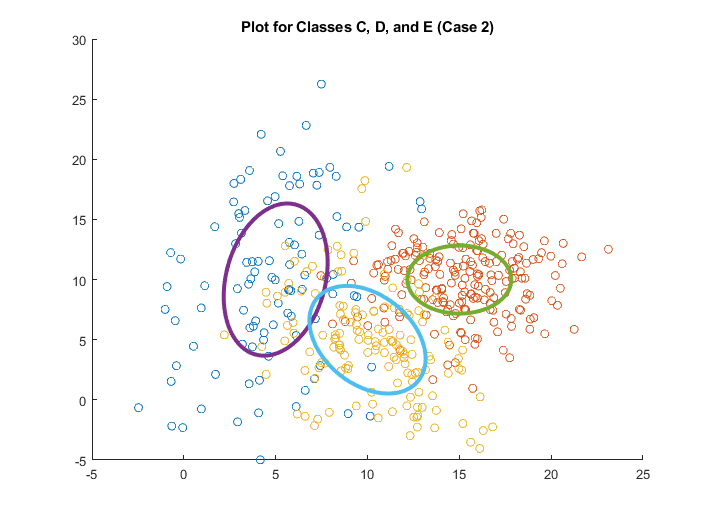
**4A Department of Systems Design Engineering**

**Due Date: Friday, February 12, 2019**

1. **Generating Clusters**

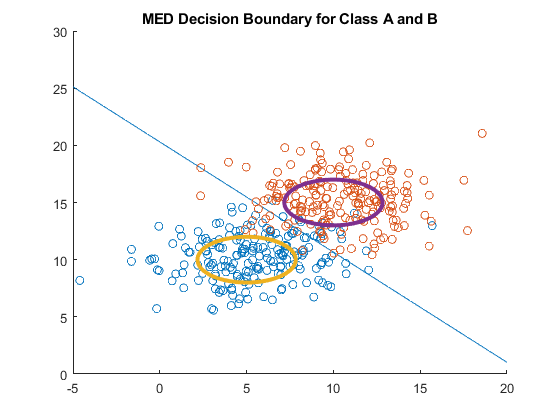
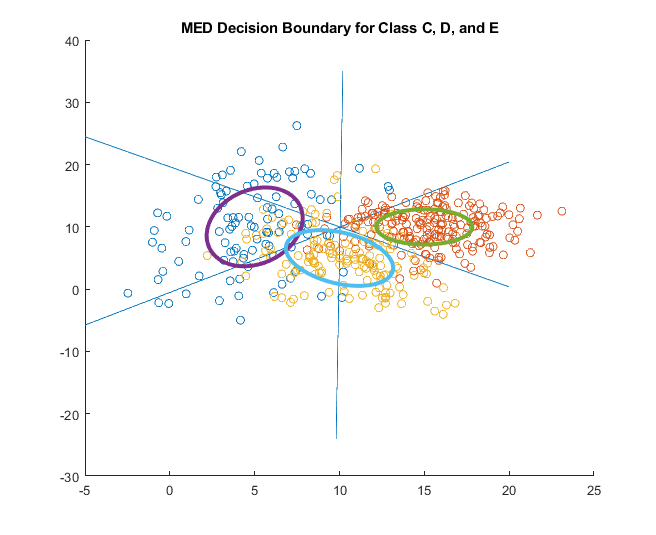
The graphs generated as as shown below.





*Q: Visually, how does the unit contour relate to the cluster data?*

A: The unit standard deviation contour represents 1 standard deviation away from the mean of the data. The contour appears to be centered at the mean of the data and contain the majority of the central points of the data, but be about a third of the way to the farthest data points. It is shaped roughly the same as the data. This makes sense because the data was generated as a normal distribution. In a normal distribution 68% of the data falls within one standard deviation which would fall within the contour. The rest of the data is more than one standard deviation away.

1. **Classifiers  
     
   2.1 Minimum Euclidean Distance (MED)  
     
     
     
   **

*Comment on the classification boundaries.*

The MED boundary is formed by the features the same distance away from both means. On the graphs this is represented by a line separating the two or three classes from each other midway between.

*How do the different boundaries compare?*

1. **2.2 Generalized Euclidean Distance (GED)  
   2.3 Maximum A Posterioi (MAP)  
   2.4 Nearest neighbor (NN)  
   2.5 K-Nearest neighbor (kNN)**
2. **Error analysis**