**ChenZH: Visualization and Classification on Finance Data**

Summary of the report: This project aimed to perform dimensionality reduction followed by classification of a finance data set. PCA, SPCA and MDS, ISOMAP, LLE and TSNE were applied to the data set and the results were visually compared. Data transformed by PCA, ISOMAP and LLE were then passed through a logistic regression or random forest classifier and the results were compared.

Describe the strengths of the report:

This report provided detailed descriptions of both data visualization and classification techniques. The introduction and explanation of the data set is clear and relates to the bigger picture.

Describe the weaknesses of the report:

There are some technical questions to be addressed.

Evaluation on clarity and quality of writing: 4

The flow of the report is easy to follow and the content is mostly conveyed clearly. Please include line numbers next time for easier review process.

Page 1- Introduction: When describing methodology, use past tense. E.g. PCA, SPCA, MDS were applied…

Page 1- Introduction: …is to visualize the high dimensional data of “a” two-dimensional plane

Evaluation on technical quality: 4

It would be helpful to elaborate more on the purpose of the classification procedure. What is the purpose of a model that can tell us what the class of a stock is (I assume that this information is readily and easily accessible)?

Figure 7a and Figure 7f have a questionable training accuracy of 100%. I wonder if this a technical error in the code?

Overall rating: 4

Confidence on your assessment: 3