

CISC5352 Artificial Intelligence in Fintech Quiz (2)¹

¹Please turn in your workable codes and corresponding running results.

Discover HFT trading data (50 points)

- Add new variables to the four HFT datasets and check their visualization under PCA and t-SNE
- Find the outliers of the HFT data in PCA and check their original samples
- Find five nearest neighbors of each sample in the input space and its projection in the PCA subspace. How many percents of a neighbor will repeat themselves in the PCA subspace?
- Implement SVD and use it to find the important samples in each HFT data (this is a little bit challenging)

What should you turn in?

- 1. A folder that contains
 - A ppt to show details of your analytics (at MOST 30 pages)
 - your data
 - source files
 - corresponding related output.
- 2. Send the zipped file (.zip instead of ,rar) of your folder to Blackboard before the DUE