FIM 590-003 Machine Learning in Finance

Project 02. Stock classification with K-means and KNN

Due time: 11:59pm on Sunday, Oct. 27.

You can form the same group as with Project 1, or a different group. Each group can have up to four group members.

Put all group members on your project reports. For grading purposes, each student needs to submit a project report (same for all students in one group).

Publicly traded companies are usually grouped under small cap, mid-cap, and large cap companies with their growth dimension as value, core, and growth stocks.

1. Using market capitalization and price per earnings growth as metrics run clustering using K-means algorithm on companies traded in Russell 2000 index with 9 clusters. Compare the cluster positioning of your favorite companies with that of the Morningstar rating for those companies. Plot these clusters and see how these boundaries are matching with your expectations.
2. Assign labels to the clusters in the companies with their growth and capitalization based on K-means analysis or Morningstar analysis. Select a new company stock that you may be interested in and classify the stock using KNN technique.

Remarks

- You may use Yahoo Finance or Bloomberg Terminals to obtain the data.
- Data collection may take some time.
- See page 76 of the Textbook 1 for more information.
- Provide your Python code and results in your report.