

# CS5803 – Advanced Data Mining

## Group Project – Semester 1

Group 09 Team Members:

- 239155F - Dilhan M.W.S
- 239173H - Samarasinghe T.A.A.
- 239187E - Wickramasinghe J.A.D.L
- 239162A - Wijesena C.M.C.H

Dataset: CSE - Colombo Stock Exchange dataset from 1990 - 2020

Project Ideas:

1. Stock prediction
2. Portfolio optimization
3. Stock clustering
4. Market sentiment analysis
5. Anomaly detection
6. Risk analysis
7. Stock ranking
8. Stock return prediction

Data set links-

- <https://www.kaggle.com/datasets/paultimothymooney/stock-market-data><https://www.kaggle.com/code/faressayah/stock-market-analysis-prediction-using-lstm>
- <https://www.kaggle.com/datasets/borismarjanovic/price-volume-data-for-all-us-stocks-etfs>
- <https://www.kaggle.com/datasets/jacksoncrow/stock-market-dataset>
- <https://www.kaggle.com/datasets/kazanov/sentiment140> - Dataset with 1.6 million tweets

