Course Projects for FRE-GY-9733 Machine Learning for Finance

Autumn 2017

1. Analysis of banks performance and business models

- **Data:** The FDIC call report data
- **Objective:** Identify banks' business models using data-driven Machine Learning approaches. Analyze profits for each identified group of banks.
- **Literature:**
  - R. Roengpitya et. al. "Bank Business Models" (2014) http://www.bis.org/publ/qtrpdf/r_qt1412g.pdf
  - M. Fethi and F. Pasiouras, "Assessing Bank Performance with Operational Research and Artificial Intelligence Techniques: A Survey", European Journal of Operation Research, 204(2), 2010, pp.189-198.
  - R. Ayadi and W. de Groen, "Banking Business Models Monitor 2014 Europe" https://www.ceps.eu/system/files/Banking%20Business%20Models%202014.pdf

2. Prediction of stock returns and portfolio selection using fundamental analysis

- **Data:** The WRDS fundamentals data
- **Objective:** Explore models that use stocks' fundamentals to predict future returns (6M and 1Y) and select an investment portfolios
- **Methods:** Linear Regression, Feedforward Neural Networks, Recurrent Neural Networks, plus whatever else you want to try (CART, SVM, etc.)
- **Literature:**
  - J. Piotroski, "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers", Journal of Accounting Research, 38 (2000), p.1
  - J. Ararbannell and B. Bushee, "Abnormal Returns to a Fundamental Analysis Strategy", The Accounting Review, v.73(1) 1998, pp. 19-45.
  - E. Fama and K. French, "Profitability, Investment and Average Returns", Journal of Financial Economics 82(3) 2006, pp. 491-518.

- Kao-Yi Shen, "Implementing Value Investing Strategy by Artificial Neural Network", International Journal of Business and Information Technology, 1 (2011), p.12
- B. Vanstone et. al. "Stockmarket trading using Fundamental Variables and Neural Networks", Bond Business School (2010), http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1156&context=infotech_pubs
- B. Yildiz and A. Yezegel, "Fundamental Analysis with Artificial Neural Network", The International Journal of Business and Finance Research, v.4 (1), 2010, pp. 149-158, http://www.theibfr2.com/RePEc/ibf/ijbfre/ijbfr-v4n1-2010/IJBFR-V4N1-2010-10.pdf
- A. Basu and A. Ashwood, "The Quest for Alpha: Can Artificial Neural Networks Help?", JASSA, 2014, March (1), pp. 13-18.
- "Deep Learning and Long-Term Investing", parts 1, 2, 3, at http://www.euclidean.com/deep-learning-long-term-investing-1/.

3. Machine Learning for Corporate Finance (making an AI CFO)

- **Data:** The WRDS fundamentals data
- **Objectives:**
  - Use Reinforcement Learning approach to build an empirical model of optimal corporate capital structure. Reformulate the previous econometric and dynamic programming models as a problem of RL.
  - Use Inverse Reinforcement Learning to learn the corporate objective function, as implied by its investment, financing and dividend decisions.
- **Literature:**
  - M.J. Flannery and K. W. Hankins, "Estimating Dynamic Panel Models in Corporate Finance", Journal of Corporate Finance, vol. 19 (2013), pp. 1-19.
  - V. A. Dang, M. Kim, and Y. Shin, "In Search of Robust Methods for Dynamic Panel Data Models in Empirical Corporate Finance", Journal of Banking and Finance, v. 53 (2015), pp. 84-98.

- I. Strebulaev and T.M. Whited, "Dynamic Models and Structural Estimation in Corporate Finance", Foundation and Trends in Finance, Vol.6, Nos.1-2 (2011), 1-163.
- M.C. Jensen, "Value Maximization, Stakeholder Theory, and the Corporate Objective Function", Journal of Applied Corporate Finance, 14(3) 2001, pp. 8-21.
- A. Damadaran, "The Objective in Corporate Finance", http://people.stern.nyu.edu/adamodar/pdfiles/acf3E/presentations/mgtobj.pdf
- https://www.youtube.com/watch?v=RZYuuX6FOg
- https://www.youtube.com/watch?v=5TeThVJFyPU
- D. Amaya et. al., "Dynamic Risk Management: Investment, Capital Structure, and Hedging in the Presence of Financial Frictions" (2012), http://publications.ut-capitole.fr/15366/1/wp_tse_330.pdf

4. Optimal Capital Structure for Commercial Banks.

- **Data:** The FDIC Call Report Data
- **Objective:** Address the problem of optimal capital structure of a commercial bank as a stochastic optimal control problem, and solve it using Reinforcement Learning and Neural Networks.
- **Note:** You do not necessarily have to follow the particular computational algorithm described in this paper, this would be a subject for a separate discussion.
- **Literature:**
  - C. Braaten et. al. "Optimal Capital Structure in Depository Financial Institutions", Norwegian University of Science and Technology (2015).
  - T. Kochubey and D. Kowalczyk, "The Relationship between Capital, Liquidity and Risk in Commercial Banks".

5. Deep Portfolio Theory – analyze the ability of a Neural Network based approach to uncover market factors.

- **Data:** synthetic data and S&P500 daily returns data

- **Objective:** Compare the PCA and Neural Networks in their ability to uncover market factors, and apply to portfolio construction.
    - Generate synthetic stock return data with one market factor and auto-regressive residuals (as e.g. in the paper by J. Yeo and G. Papanicolaou, https://arxiv.org/pdf/1611.05571.pdf). Then use the PCA to uncover and estimate mean reversion and volatility of residuals.
    - Try to uncover the factor(s) and the residuals using the Deep Learning approach (the Deep Portfolio Theory by Heaton et. al.
    - Do the analysis with the S&P500 data. You may want to add technical indicators as additional predictors, similar to Troiano et. al.
- **Literature:**
    - J. Yeo and G. Papanicolaou, "Random Matrix Approach to Estimation of High-Dimensional Factor Models" (2016) (https://arxiv.org/pdf/1611.05571.pdf)
    - J. Heaton, N. Polson, J. Witte, "Deep Portfolio Theory" https://arxiv.org/abs/1605.07230
    - N. Polson and V. Sokolov, "Deep Learning: a Bayesian Perspective" (2017) https://arxiv.org/abs/1706.00473
    - L.Troiano, E. Mejuto, and P. Kriplani, "On Feature Reduction using Deep Learning for Trend Prediction in Finance" (2016) https://arxiv.org/abs/1704.03205

6. Analysis of causal relations and information propagation between stocks.

- **Data:** S&P500 intraday stock data
- **Objective:** Build a model of causal relationship between stock price changes. Use the results to cluster stocks in groups which would have their leaders and laggards.
- **Literature:**
    - V. Boginski, S. Butenko, and P. M. Pardalos, "Statistical analysis of financial networks", Computational statistics & data analysis, 48(2):431–443, 2005.
    - O. Kwon and J.S. Yang, "Information Flow Between Composite Stock Index and Individual Stocks" (2007), https://arxiv.org/abs/0708.0063

- K.T. Chi, J. Liu, and F. CM Lau, "A network perspective of the stock market, Journal of Empirical Finance, 17(4):659–667, 2010.
- R. B. Roy and U. K. Sarkar, "Identifying influential stock indices from global stock markets: A social network analysis approach", Procedia Computer Science, 5:442–449, 2011.
- D. Dindi, A. Ozturk, and K. Wyngarden, "Predicting Stock Movements Using Market Correlation Networks" https://pdfs.semanticscholar.org/f9ef/bdc05d1814c29afe52c0fb8a3de1c9dedf0e.pdf
- L. Sandoval, "Structure of a Global Network of Financial Companies based on Transfer Entropy", Entropy 2014, 16, pp. 4443-4482.

7. Portfolio optimization with regime changes

- **Data:** S&P500 stocks, plus macro-variables (from Bloomberg?)
- **Objective:** Investigate portfolio construction under time-varying economic regimes.
- **Literature:**
  - C. Turner and J. Han, "Portfolio Optimization under Time-Varying Economic Regimes" http://cs229.stanford.edu/proj2009/HanTurner.pdf, and references therein.

8. Index tracking using Reinforcement Learning and/or other techniques

- **Data:** S&P daily stock prices
- **Objective:** Construct a stock portfolio with a small number of stocks that mimics or closely tracks the returns of S&P500 index.
- **Literature:**
  - J. Park, D. Yang, and K. Park, "Approximate Dynamic Programming-Based Dynamic Portfolio Optimization for Constrained Index Tracking", International Journal of Fuzzy Logic and Intelligent Systems, vol, 13, no.1 (2013), pp. 19-30, and references therein.
  - T. Roncalli and G. Weisang, "Tracking Problems, Hedge Fund Replication and Alternative Beta", (2008), http://www.thierry-roncalli.com/download/particle-filter.pdf

Groups of Students

Group Number	Student Name	NYU Net ID	Project Number and Title
1	Haoran Su (Leader) Ruijing Yang Wanting Chen Bokai Xiang	hs3265 ry829 wc1418 bx360	5. Deep Portfolio Theory – analyze the ability of a Neural Network based approach to uncover market factors.
2	Srikanth Myskar (Leader) Smrati Hans Ridhish Mukund Batavia Sreecharanreddy Pothireddi	sm7039 sh4494 rmb589 sp4511	8. Index tracking using Reinforcement Learning and/or other techniques
3	Haotian Cai (Leader) Jason Ellis Chen Yang Yutian Ding	hc1870 je1317 cy1133 yd983	2. Prediction of stock returns and portfolio selection using fundamental analysis
4	Zhijie Gao (Leader) Zhitong Ye Yifeng Jin Ao Yin	zg682 zy924 yj1088 ay1191	6. Analysis of causal relations and information propagation between stocks
5	Xiao Li (Leader) Huiqi Tian Ruiqing Yu Shangli Xu	xl1947 ht980 ry828 sx616	7. Portfolio optimization with regime changes Data
6	Bohui Xi (Leader) Yuhan Liu Tianrui Zhao	bx363 yl4386 tz962	Optional Topic.

7	Ran Xu (Leader) Zhao Ma Qiliang He	rx359 zm746 qh446	2. Prediction of stock returns and portfolio selection using fundamental analysis
8	Ming Gong (Leader) Tianyu Xu Chang Liu Yuehan Liu	mg5313 tx401 cl4094 yl4478	8. Index tracking using Reinforcement Learning and/or other techniques