FRE 7773
Machine Learning for Finance
Fall 2022
Tuesdays 5:30-8:00p

Instructor: Prof. Sandeep Jain
Email: sj1612@nyu.edu
Office Location/Hours: By appointment

Course Description:
This course is an introduction to Machine Learning concepts and its application to the financial industry. The purpose is to become knowledgeable enough about the field to pursue individual topics further. In addition, the topics will include practical implementation of the techniques in Python on financial data or other sample data when financial data not available.

Requirements: The class will use Jupyter Notebooks and all students are expected to run code and do homeworks in that environment.

Texts: Not required for purchase. Throughout the course, each lecture will draw on material from the textbooks below, which are excellent resources at different levels of difficulty.

- Pattern Recognition and Machine Learning, Bishop
- The Elements of Statistical Learning by Hastie et al.
- Deep Learning by Goodfellow et al.
- Neural Networks and Deep Learning, Michael Nielsen
- [JVP] Python Data Science Handbook by VanderPlas
- [WM] Python for Data Analysis by Wes McKinney
- [AG] Hands-On Machine Learning with Scikit Learn and TensorFlow by A. Geron.
- [SB] Reinforcement Learning: An Introduction by Sutton and Barto

Grading: 10 HWs (70%), final project (30%).

Course Outline: Material will not necessarily be taught in this order and some topics are subject to be added/removed.

Part I: *Classical Machine Learning in Finance*

- Introduction to machine learning in finance and Python
- End to End Machine Learning using Scikit-Learn
- Mathematical Foundations of Machine Learning
- Supervised Learning (with applications in Finance):
    a. kNN
    b. Regression

    c. Kalman filtering
    d. Classification and Categorization
    e. Support Vector Machines
    f. Decision Trees and Random Forests
- Unsupervised Learning: Agglomerative clustering for portfolio construction
- Feature Selection: Constrained regression and VaR hedging
- Classification with Imbalanced data sets

Part II: *Neural Networks and Deep Learning*

- Working with Tensor Flow and Keras
- Artificial Neural Networks
- Training Deep Neural Networks (Vanishing/exploding gradients, Regularization Techniques, fine tuning hyperparameters)
- Convolution Neural Networks
- Recurrent Networks (GRU, LSTM, Deep RNN, Natural Language Processing)
- AutoEncoders and GANs

Reinforcement Learning (time permitting)