SA.310.817.51: Machine Learning for Finance with Python

S. Yanki Kalfa

Intersession, 2023

E-mail: Yanki.Kalfa@rady.ucsd.edu
Office Hours: By Appointment
Office: Zoom

Web: **Teaching Material**
Class Hours: 1pm-4pm
Class Room: Zoom

Course Description

The emergence of freely available data lead to widespread application of machine learning (ML) in finance. Machine learning methods are used in many areas of finance such as fraud detection, credit card applications, portfolio construction, and stock return prediction.

This course will start with the basics of financial forecasting and build up to the cutting edge of machine learning. We will overview the basic process of data science and machine learning techniques for finance. This course will focus on the tools necessary to forecast stock returns and portfolio construction.

The ability of learning patterns from data and making accurate predictions with new data is what makes ML a powerful tool for financial forecasting and decision making. This course will touch on the theory of ML algorithms but not dive into the technical details. We will instead focus on the applications of ML.

Hands-on applications are an important part of this course. Students will be introduced to some theoretical properties of ML algorithms and immediately will be shown how to implement them. The purpose of this approach is to train the students to develop ML projects independently and confidently.

Prerequisites

Basic knowledge of probability and econometrics. Familiarity with linear algebra. We will use Google Colab for our hands-on sessions and assignments. Please make sure you have a Google account.

Course Objectives

1. Basic Machine Learning process
2. Python and Jupyter notebooks
3. Manipulation of financial data
4. Application of Machine Learning algorithms to financial data

Course Structure

Each class meeting will be organized in the following way. We will first go over some theoretical properties of machine learning algorithms, then we will apply the algorithm in real time, and finally students will go over some Machine Learning tasks during class.

Resources

We will use three textbooks as fundamental material for this class. You are not required to buy these textbooks. The first two texts listed are free to download.

- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning. Vol. 112. New York: springer, 2013. (ISLR)
- Hastie, Trevor, Robert Tibshirani, Jerome H. Friedman, and Jerome H. Friedman. The elements of statistical learning: data mining, inference, and prediction. Vol. 2. New York: springer, 2009. (ESL)
- Elliott, Graham, and Allan Timmermann, Economic Forecasting. Princeton University Press, 2016. (EF)

The sources listed above are great fundamental readings. EF textbook is probably the most up to date forecasting textbook that takes a decision science approach to forecasting. ISLR is a great introduction to Machine Learning. ESL is a little bit more advanced and technical relative to ISLR but cover the same topics and more.

Assignment

You will have 3 in-class assignments. Assignments are equally weighted.

Grading Policy

We will adhere to the SAIS grading policy.

- A: 94 or above
- A-: $[90, 94)$
- B+: $[87, 90)$
- B: $[84, 87)$
- B-: $[80, 84)$
- C: Below 80

Schedule

The schedule is tentative and subject to change. The learning goals below should be viewed as the key concepts you should grasp after each lecture.

Lecture 1: Introduction to Class

- Intro to class
- Overview of ML
- Differences between Supervised and Unsupervised Learning
- Data science terminology
- Standard Time Series forecasting models

Readings:

- EF: Chapters 2, 7, 15, 16
- ISLR: Chapter 2

Lecture 2: Multivariate Regression and Model Selection

- ARMAx
- Loss Functions
- Lasso
- Ridge
- Elastic Net

Sources/Readings:

- EF: Chapters 6
- ISLR: Chapter 6
- ESL: Chapter 3

Lecture 3: Decision Trees and Random Forests

- CART trees
- Random Forests
- Boosted Trees
- Manual Hyperparameter Search

Sources/Readings:

- ESL: Chapters 9, 11, 15

- ISLR: Chapter 8

Lecture 4: Automatic Hyperparameter Search and Evaluation

- Cross Validation
- Model Evaluation
- Forecast Evaluation
- Out of Sample R^2

Sources/Readings:

- ESL: Chapter 7
- ISLR: Chapter 5
- RF: Chapter 15, 17

Lecture 5: Intro to Deep Learning with Tensorflow and Keras

- Feed Forward Neural Nets
- Activation Functions
- Loss Functions
- Backward Propagation
- Optimization

Sources/Readings:

- ESL: Chapter 11
- ISLR: Chapter 10