FNCE 5352 – Financial Programming and Modeling
January 17, 2023

1. Course Description:

This course will introduce the students to a wide variety of algorithms that are used in machine learning applications. Students will code a few algorithms completely and learn to use software packages that implement others. Instruction and assignments will use the R and Python programming ecosystems. Students will be exposed to Machine Learning at scale using the Keras and TensorFlow libraries. Throughout the course, special attention will be given to applications of these algorithms to finance.

2. Instructors

Ed Hayes: ehayes@uconn.edu
Matt McDonald: matthew.mcdonald@uconn.edu

Chi Zhang (TA): chi.13.zhang@uconn.edu

3. Course Delivery

14 lectures

4. Academic Integrity

Students must adhere to the University of Connecticut Student Code, which can be found at: https://community.uconn.edu/the-student-code-pdf/. Assignments and/or quizzes must be completed individually.

5. Required Texts

Part 1 (R and RStudio, with Professor McDonald)

We will be using the book “R for Data Science” by Hadley Wickham and Garrett Grolemund. The book is available online at https://r4ds.had.co.nz/index.html. It is free and licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0

If you’d like a hard copy of the book, it is available from Amazon (https://www.amazon.com/Data-Science-Transform-Visualize-Model/dp/1491910399/ref=sr_1_3?ie=UTF8&qid=1548809834&sr=8-3&keywords=r+for+data+science)

Additionally, we will use the book “Feature Engineering and Selection” by Max Kuhn and Kjell Johnson. Physical copies are sold by **Amazon** and **Taylor & Francis**. An online version is available at https://bookdown.org/max/FES/

Part 2 (Python, with Professor Hayes)

We will be using selections from the first 2 books listed here. These books can be read using the O'Reilly website, which also comes as an app that you can use. You can read while connected and you can also download many titles. This includes O'Reilly titles, Manning titles, and a variety of others. This is a great resource!

1) Tatsat, Hariom, *et al*. *Machine Learning and Data Science Blueprints for Finance: from Building Trading Strategies to Robo-Advisors Using Python*. O'Reilly Media, 2020. **("Blueprints")**

   We are going to like working with this book! The Python environment is relatively simple to set up, "holes" in the code base are easy enough to fix, and the spread of topics is good.

2) Tony Guida, *et al*. *Big Data and Machine Learning in Quantitative Investment*. Wiley, 2019. **("BigDataML")**

   We will cover Chapter 12 of this book. I will give you notes, so you do not have to buy it. It's a good book, but book #3 is a better value if you only want to buy 1.

3) Stefan Jansen, *Machine Learning for Algorithmic Trading: Predictive models to extract signals from market and alternative data for systematic trading strategies with Python, 2nd Edition*. Packt Publishing, 2021

This was the primary text last year, but I moved it to secondary this year. It is an excellent in-depth survey of many advanced topics in ML with compelling applications to finance. It is over 1500 pages long! It is well worth your attention, but I think our 7-8 weeks together are better served by book #1.

4) Géron, Aurélien. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly, 2020.

This used to be the primary text for the Python section of the class. Chapter 2 has a particularly excellent end-to-end example of how the ML process works in practice using the familiar example of Linear Regression as it subject.

6. Part 1 (R and RStudio) Prerequisites

Course Materials

Course materials can be found at https://github.com/mattmcd71/fnce5352_spring2023

This repository will be updated throughout the semester.

R Fundamentals

The course assumes some understanding of the R programming language. If you would like to get a basic introduction to the R programming language, please visit the following link:

https://www.rstudio.com/online-learning/

R Installation

We will be periodically using R and RStudio interactively during the class instruction. If you would like to follow along during the class, please follow these instructions

Local Installation Instructions:

R

We'll be using the most recent version of R locally but I believe that anything > 3.5.1 should be fine.

R can be downloaded from the following link: https://www.r-project.org/

RStudio

RStudio is an Interactive Development Environment for the R programming language. It is very useful. You can download it at:

https://www.rstudio.com/products/rstudio/download/

7. Part 2 (Python) Prerequisites

If you have time, start reading the Blueprints book and try to complete some of the setup.

University of Connecticut | Master of Science in Financial Risk Management

Lessons and Assignments

Lecture Date	Topic	Assignment	Reading assignment (before next class)
17-Jan	Intro to R and RStudio		R4DS: Sections 1-6 Articles: "Good enough practices in scientific computing" "Excuse me, do you have a moment to talk about version control?"
24-Jan	Analytic Workflow & Visualization	R4DS: 3.2.4 (Page 6): Exercise 5 3.3.1 (Page 12): Exercise 2 3.6.1 (Page 20): Exercise 1 4.4: Practice 3 (Page 41) (Due 1/31)	R4DS: 7-10 FES: Chapters 1 & 3
31-Jan	Data Wrangling with the Tidyverse	R4DS: 5.2.4 (page 49): Exercises 1, 3 5.3.1 (page 51): Exercise 1 5.5.2 (page 59): Exercises 2, 5 (Due 2/7)	R4DS: Sections 11-13
7-Feb	Modeling – Introduction, Data Usage, & Feature Engineering		R4DS: Sections 14-19
14-Feb	Modeling – Resampling & Grid Search		R4DS: Sections 20-25
21-Feb	Regression in R		
28-Feb	Classification in R	Credit Modeling Project (due end of semester)	

PART 2			
	Topics	**Case Studies**	**Chapters**
7-Mar	Machine Learning in Finance: Modeling Perspective and Framework; Regression and Classification	Stock Price Prediction; Yield Curve Prediction	Blueprints 1, 2, 4, 5
21-Mar	Decision Trees, Random Forests, and Gradient Boosting	Fraud Detection; Default Prediction	Blueprint 6
28-Mar	Neural Networks, Time Series Modeling, LSTM	Stock Price Prediction; Bitcoin Trading	Blueprints 3
4-Apr	Unsupervised Learning: PCA and Kernel PCA	Eigen Portfolios; Yield Curve Modeling; Bitcoin Trading	Blueprints 7
11-Apr	Unsupervised Learning: t-SNE; Clustering Techniques	Pairs Trading; Hierarchical Risk Parity	Blueprints 8
18-Apr	Reinforcement Learning	RL-based Trading Strategy	BigDataML 12
25-Apr	Reinforcement Learning	Portfolio Allocation	Blueprints 9