**IE517 Machine Learning in Finance (Fall 2024) – CRN: 74048**

Course Website (Canvas): <https://courses.illinois.edu/schedule/2024/fall/IE/517>

Instructor: Matthew Murphy (mdmurph@illinois.edu)

Location: 2101 Everitt Lab

Time: Tuesday/Thursday 9:30-10:50 am (Office hours: 1-3:00 pm T/Th, 3254 DCL, or by appt)

**Course Outline**

Machine Learning includes the design and the study of algorithms that can learn from experience, improve their performance, and make predictions. This course is designed specifically and exclusively for MSFE first semester students. It features rigorous coding exercises in python and acts as preparation for later courses. Students will learn the concepts behind different supervised machine learning algorithms and implement them in Python using advanced packages; pandas, NumPy, and scikit-learn. All the data for this course features unique real-world financial datasets produced through interactions with MSFE practicum students and sponsors.

**Course Objectives**

**Course Learning Outcomes**

Through the completion of this course students are expected to develop competencies in:

* Basic machine learning concepts and models. An intuitive, rather than mathematical understanding, and focus on real-world interpretation and applicability of models. Become conversant with machine learning.
* Data analytics project framework. How to begin, implement and complete a data analytics project using machine learning tools. Data acquisition and preprocessing, to feature extraction to model selection and feature engineering to model fitting/evaluation and improvement using ensemble techniques and hyperparameter tuning. Learn how to best ask questions of the data to derive meaningful insights.
* Python implementation environment. Ability to perform a complete data analytics project, including all phases, competently in python coding. Understanding of the Scikit Learn package of ML models and use SKLearn Documentation to answer basic technical/implementation questions.
* Exploratory data analysis and visualizations of datasets of a comparable scale to MSFE practicums.
* Thinking independently and experimentally. Proceeding without explicit directions to solve data and coding problems. Develop comfortability with an ad hoc, trial-and-error experimental approach.
* Working as part of a distributed team. Coordinating to utilize specialization, division of labor and economy of scale and to develop effective behind-the-scenes teamwork and communications skills. Foster creativity, collaboration, curiosity and innovative problem solving.

**Textbook and other materials**

Main (required) textbook:

* Raschka, Sebastian and Yuxi Hayden Liu and V. Mirjalili. 2022. *Machine Learning with PyTorch and Scikit-Learn*. Packt Publishing. Available free from Safari Books Online with your valid student email address: (https://www.safaribooksonline.com/library/view/python-machine-learning/9781787125933/ )



* A Github repository for the code and data provided in the text is available at: <https://github.com/rasbt/machine-learning-book>
* Balaraman, G. & Ballabio, L. *QuantLib Python Cookbook. Hands-on Quantitative Finance in Python.* 2022. LeanPub. Available from: <http://leanpub.com/quantlibpythoncookbook>

Other course materials may be obtained from the following:

* Bowles, Michael. *Machine Learning in Python*. 2015. Wiley. Available free online from Wiley Online Library: (https://onlinelibrary.wiley.com/doi/book/10.1002/9781119183600 )
* Hackeling, Gavin. *Mastering Machine Learning with sci-kit learn*. 2017. Second Edition. Packt Publishing. Available free from Safari Books Online with your valid student email address: (https://www.safaribooksonline.com/library/view/mastering-machine-learning/9781788299879/ )
* Alpaydin, Ethem. *Introduction to Machine Learning*. 2014. The MIT Press. Available free online from IEEE Xplore: (<https://ieeexplore-ieee-org.proxy2.library.illinois.edu/xpl/bkabstractplus.jsp?bkn=6895440&SID=EBSCO:edseee>)
* Moncecchi, Guillermo and Raúl Garreta. 2013. *Learning scikit-learn: Machine Learning in Python*. Packt Publishing. Available online free from Safari Books Onlline: (<https://www.safaribooksonline.com/library/view/learning-scikit-learn-machine/9781783281930/> )
* Scikit-learn user guide. Release 1.5.1. Available at: (<https://scikit-learn.org/stable/user_guide.html> )

Other assigned or optional readings will be made available via the course Compass site. Lecture notes, homework assignments, and other materials will be posted on the course Compass site.

**Course Topic Outline**

Module 1 - Python for Finance (pandas, numpy, and producing efficient code)
Module 2 – Machine Learning methods and Data Analytics

Module 3 – Machine Learning for Option Pricing

Module 4 – Deep Learning for Advanced Option Pricing

**Course requirements**

Grades will be determined as follows:

Individual homework assignments (4 at 10% each) 40%

Group project 40%

Participation (attendance/engagement) 10%

Final Evaluation 10%

**The maximum group size is four persons.**

The Final Evaluation (which will be the “final” exam for this 8 week course) is scheduled for **8-11 am Tuesday, December 10, 2024**. (Last day of instruction).

**DataCamp for the Classroom**DataCamp is a subscription, web-based learning platform for a variety of data and programming skillsets. As part of the DataCamp for the Classroom program, you will have free access to the entire DataCamp course library for the duration of this class, plus a free trial for six months. Ordinarily this is a monthly subscription cost. I will create an assignment list of chapters and courses that align with the modules as we cover them. You may proceed at your own pace and feel free to explore beyond the confines of the assigned modules.

**Office hours**

Tuesday/Thursday from 1:00 to 3:00 pm (3254 DCL) or by appt. (Zoom or MSTeams is also an option).

**Attendance**

There is no specific attendance requirement, although it is assumed that regular attendance will improve your performance on the assignments. Because this course is only half a semester, missing even one class meeting will put you at a disadvantage. You are expected to attend all regularly scheduled classes, but you will not be penalized until you have more than two unexcused absences. A portion of your grade will be determined by Participation which includes attendance and your engagement in the class.

**Modality**

This course is offered in only a traditional physical classroom. 2x week. There are only 13 scheduled class meetings for this half semester course. Missing more than 2 sessions without an excused absence will result in the reduction of one letter grade. There will be no recordings available for the lab sessions.

**Course Management System**

Illinois Canvas is the communication and course material access system used in this class. You may access the course website using an Internet browser at <https://canvas.illinois.edu/>

You will need to login using your NetID and your NetID password. Course information and materials, as well as course grades, will be available via the Illinois Canvas website. You are expected to access the site regularly. If you are not familiar with Canvas, please view “Getting Started with Canvas” located at <https://movetocanvas.web.illinois.edu/getting-started/>

**Academic integrity**

Please be aware of the University’s policies regarding academic integrity. See

<https://studentcode.illinois.edu/article1/part4/1-401/>