Rady | **School of Management**
UNIVERSITY OF CALIFORNIA SAN DIEGO

MGTF 495 SPECIAL TOPICS:
DATA SCIENCE FOR FINANCE USING PYTHON
Spring 2022, Master of Finance

PROFESSOR VOLKAN VURAL
EMAIL: vvural@ucsd.edu
OFFICE HOURS: By appointment
TEACHING ASSISTANT:
Yanki Kalfa (skalfa@ucsd.edu)
Sri Pamidi (spamidi@ucsd.edu)
Mengjie Wang (mew006@ucsd.edu)

COURSE OBJECTIVES

Data science is an emerging field that brings together vast amounts of data with techniques in a range of areas including data management, statistical analysis and machine learning.

This course, a part of the Rady Master of Finance Program, will introduce you to a collection of powerful, open-source, tools needed to manage and analyze finance data, and to conduct data science in finance applications. We will overview the basic process of data science and machine learning techniques useful for financial data analysis.

The ability of learning patterns from data and making accurate predictions on new instances makes Machine Learning a powerful tool for Business Intelligence and Financial Forecasting since it helps us transform the raw data into better decisions. This course will go over the theory of Machine Learning algorithms but not dive into the technical details. We will rather focus on how these algorithms work and how we can apply them to financial applications.

Hands-on applications are an important part of this class. Not only will the students learn the theory of data science but also apply this theory to real world applications during the class time and after the class for their homework and project. Our goal with this hands-on approach is to train the students so that they will be able to develop data science projects independently with confidence by the end of the quarter. Popular Python based tools will be used in analytical components of the class and in practical exercises

COURSE PREQUISITES

1. Ability to write programs in Python: functions, control structures, string handling, arrays and dictionaries
2. Familiarity with basic probability. To brush up on this, you might look at the following book, available online: Grinstead and Snell. Introduction to Probability
3. Familiarity with basic linear algebra. To brush up on this, we recommend: Gilbert Strang. Linear algebra and its applications.

LEARNING OBJECTIVES
- Basic process of data science
- Python and Jupyter notebooks
- An applied understanding of how to manipulate and analyze financial datasets
- Basic statistical analysis and machine learning methods

- Interpret descriptive statistics such as mean, variance, mode, median, coefficient of
- variation, outliers, kurtosis
- Understand and apply various machine learning algorithms for classification, clustering
- and regression
- How to effectively visualize results

COURSE MATERIALS

There are no required text for the course. Some optional texts are:
- Trevor Hastie, Robert Tibshirani, and Jerome Friedman, The elements of statistical learning (2nd edition). This is available as an e-book through the UCSD library website.
- Kevin Murphy, Machine learning: a probabilistic perspective.
- Richard Duda, Peter Hart, and David Stork, Pattern classification (2nd edition).
- Stephen Boyd and Lieven Vandenberghe, Convex optimization. This is available online from Boyd's website.

COURSE GRADES

Course grades will be based on the following weighted components:
- Assignments (60%)
- Final Project (40%)

TEACHING ASSISTANTS

The teaching assistants for this course are Yanki Kalfa, Sri Pamidi and Mengjie Wang. They will be available to answer questions in person and via e-mail during the week. Their emails are included above.

PROFESSOR CONTACT INFORMATION

E-mail: vvural@ucsd.edu

STUDENTS WITH DISABILITIES

A student who has a disability or special need and requires an accommodation in order to have equal access to the classroom must register with the Office for Students with Disabilities (OSD). The OSD will determine what accommodations may be made and provide the necessary documentation to present to the faculty member.

The student must present the OSD letter of certification and OSD accommodation recommendation to the appropriate faculty member in order to initiate the request for accommodation in classes, examinations, or other academic program activities. **No accommodations can be implemented retroactively.**

Please visit the OSD Web site (osd.ucsd.edu) for further information or contact the Office for Students with Disabilities at (858) 534-4382 or fosorio@ucsd.edu.

ACADEMIC INTEGRITY

Integrity of scholarship is essential for an academic community. As members of the Rady School, we pledge ourselves to uphold the highest ethical standards. The University expects that both faculty and

students will honor this principle and in so doing protect the validity of University intellectual work. For students, this means that all academic work will be done by the individual to whom it is assigned, without unauthorized aid of any kind.

The complete UCSD Policy on Integrity of Scholarship can be viewed at: http://www-senate.ucsd.edu/manual/appendices/app2.htm#AP14

MGTF495 COURSE OUTLINE

Week 1, March 28
- Introduction to class
- Introduction to data science
- Overview of machine learning concepts and tasks
- Introduction to supervised and unsupervised learning
- Example workflow of a data science project

Week 2, April 4: Assignment 1 is released
- Decision trees
- Hands-on
- Self-practice
- Nearest neighbor classification
- Hands-on
- Self-practice

Week 3, April 11:
- Regression
- Hands-on
- Self-practice
- Evaluating Classifiers
- Hands-on
- Self-practice

Week 4, April 18: Assignment 2 is released
- Generative models
- Hands-on
- Self-practice

Week 5, April 25: Final Project is released
- Discriminative models
- Logistic regression
- Hands-on
- Self-practice
- Perceptron
- Hands-on
- Self-practice

Week 6, May 2: Assignment 3 is released
- Support vector machines
- Kernels
- Hands-on
- Self-practice
- Richer output spaces

Week 7, May 9:
- Clustering
- Informative Projections
- Hands-on
- Self-practice
- Embeddings
- Manifold Learning
- Dictionary Learning
- Hands-on
- Self-practice

Week 8, May 16:
- Deep Learning
- Hands-on
- Self-practice

Week 9, May 23:
- Deep Learning
- Hands-on
- Self-practice

Week 10, May 30:
- Project Presentations