Course	BUAN/MIS/OPRE 6341
Course Title	Applied Machine Learning
Instructor	Uri Smashnov
Term	Spring 2024
Meetings	Lecture with Virtual Attendance (Online)

Professor Contact Information

Instructor: Uri Smashnov
Email Address uri.smashnov@utdallas.edu
Office hours: virtual office hours via MS Booking: TBD

Teaching Assistant: TBD
Email:
Office hours:

Course Modality and Expectations

Instructional Mode	Lecture - recorded
Lecture Time	Lectures will be pre-recorded and class materials made available on a Module by module basis. Projects will be introduced in the synchronous session that will be recorded and posted on eLearning.
Course Platform	Course materials will be posted the eLearning page of the course.
Asynchronous Learning Guidelines	Posted on on eLearning

Course Pre-requisites, Co-requisites, and/or Other Restrictions

Please consult CourseBook and your academic advisor as of August 5th, 2023 the pre-requisites are listed as following: Prerequisites: (BUAN 6356 or BUAN 6383 or MIS 6386 or BUAN 6324 or MIS 6324 or OPRE 6399) and (OPRE 6359 or BUAN 6359).

The course is taught in Python. Students are expected to have prior programming and Python experience.

Course Description

In this advanced course, we will delve into various machine learning models tailored for business analytics. Our focus will encompass both supervised and unsupervised models, intricate non-linear regression techniques, resampling methods, and introduction to neural networks and AI-driven models essential for data analytics. Python, a leading programming language in this field, will be our primary tool. Given the course's advanced level, a proficient understanding of Python is assumed, and we will allocate minimal time to revisiting Python basics. Additionally, we will explore crucial aspects of machine learning interpretability, including global and local Explainability, model debugging, and fostering trust in model outputs.

Student Learning Objectives/Outcomes

Upon completion of this course, students will have the skills to:

1. Distinguish between supervised learning methods (such as classification) and unsupervised learning methods (like clustering).
2. Select the appropriate machine learning technique for a given dataset and business scenario.
3. Skillfully perform feature engineering to optimize machine learning models and proficiently use Python for analytical purposes.
4. Design and execute machine learning experiments aimed at minimizing overfitting and achieving the best possible results with the available data in relation to specific business challenges.
5. Grasp and implement concepts related to the interpretability of machine learning models.

Required Textbooks and Materials

Required Texts

Textbooks are not required for this course. Supporting material will be provided through eLearning course page.

Required Hardware and Software

The lecture materials are tested only with Windows 10 and latest MacOS running Intel Chip. Mac with M1/M2 chip is not tested. Some libraries might require compilation and additional steps to be installed on Mac OS M1/M2 and students should contact professor or TA if help with installation is needed.

Required Memory: at least 8 GB memory is required, 16 GB memory is recommended. Students can rent/provision Linux/Windows box on one of the Cloud providers and use for the class if personal computer doesn’t meet memory requirements.

Suggested approach: provision AWS EC2 Ubuntu box (t3a.xlarge) with Spot pricing and only "start" it while using it, rest of the time only storage charges will apply. It is possible to open tunnel to Jupyter notebook only from your PC to make the setup very secure. Neural nets module will be taught using Google Colab. Premium ($10 per month – only 1 month will be needed) subscription is advised to enable GPU use.

Software: Python 3.10.11 - Additional packages you need to have for this course are H2O-3, numpy, scipy, sklearn, pandas, matplotlib, seaborn, graphviz, TensorFlow, and keras. Exact version of Python and the libraries will be posted on eLearning.
H2O-3 has Java version 1.8 or later requirement.
Use of Python virtual environment is mandatory – details are provided in the in eLearning.

Recommended Texts

Introduction to Machine Learning with Python
ISBN 978-1-4493-6941-5
Publisher O'Reilly Media, Incorporated
https://learning.oreilly.com/library/view/introduction-to-machine/9781449369880/

Introduction to Machine Learning
Author Ethem Alpaydin and Francis Bach
ISBN 9780262043793
Publisher MIT Press

Machine Learning at Scale with H2O
Authors: Gregory Keys and David Whiting
ISBN: 978-1800566019
Online Book (Free with UTD login):
https://learning.oreilly.com/library/view/machine-learning-at/9781800566019/

Online Sklearn and H2O-3 documentation and tutorials.

Textbooks and some other bookstore materials can be ordered online or purchased at the UT Dallas Bookstore.
Some books are available online and are free with UTD login. Playlist can be found here: Oreilly Learning Books playlist

Technical Requirements

In addition to a confident level of computer and Internet literacy, certain minimum technical requirements must be met to enable a successful learning experience. Please review the important technical requirements on the on the Getting Started with eLearning webpage.

Course Access and Navigation

The course can be accessed using your UT Dallas NetID account on the eLearning website. Please see the course access and navigation section of the Getting Started with eLearning webpage for more information.

To become familiar with the eLearning tool, please see Student eLearning Tutorials webpage.

UT Dallas provides eLearning technical support 24 hours a day/7 days a week. The eLearning Support Center includes services include a toll-free telephone number for immediate assistance (1-866-588-3192), email request service, and an online chat service.

Communication

All class discussions will be performed on MS Teams under appropriate channel. For example, we will use Project channel for Project discussions. Use of email is appropriate for personal questions only.

Distance Learning Student Resources

Online students have access to resources including the McDermott Library, Academic Advising, The Office of Student Accessibility, and many others. Please see the eLearning Current Students webpage for more information.

Server Unavailability or Other Technical Difficulties

The University is committed to providing a reliable learning management system to all users. However, in the event of any unexpected server outage or any unusual technical difficulty which prevents students from completing a time sensitive assessment activity, the instructor will provide an appropriate accommodation based on the situation. Students should immediately report any problems to the instructor and also contact the eLearning Help Desk. The instructor and the eLearning Help Desk will work with the student to resolve any issues at the earliest possible time.

Student Assessments

Grading Criteria:
In line with the applied nature of this class, a large portion of the assessment will be made through homework. There will be 2 project assignments. The homework will contain some theory questions, but the majority of the material will involve implementing the different methods that we cover in class using the computer package. There will be one take-home exam. The breakdown will be:

Item	Point
Project 1	15
Project 2	25
Weekly Labs	30
Final Quiz	30
Total	100

Grading Scale:
Relative grading is a system of assessment to determine students' grades by comparing them against those of their peers. Unlike the system of absolute grading, where a student's score on a given test or assignment directly converts into a letter grade (for example, 93/100 is A, etc.), relative grading means students' marks fluctuate depending on how they did compared to others in class. Relative grading refers to a system of evaluation that allows educators to convert the outcomes of a student's test, project or assignment and adjust that final grade in relation to grades from other students in the course. Relative grading is similar to bell curving or grading on a curve, and considers the highest score as the baseline (A), relatively adjusting all others compared to that score. Student should earn a passing grade for each project and exam grading components in order to be considered for a letter grade in the range of C to A.

Note: this grading system is following the UTD/JSOM policy to keep the class grade average between B to A-.

Grade	Range of marks
A	80th percentile and above. A grade of 91 or above will guarantee an A.
A-	60th to 80th percentile.
B+	40th to 60th percentile.
B	20th to 40th percentile.
B-/C	Below 20th percentile.
F	Failing grade in any grading component. A grade of 60 or below out of 100 is a failing grade.

Accessing Grades:
Students can check their grades by clicking “My Grades” on the course menu after the grade for each assessment task is released.

Projects:
There are 2 individual projects for this course. Project details will be posted in the course.

Labs:
There are Labs as individual assessment activities.

Quiz:
There will be one Quiz at the end for this course.

- The quiz will be held during regular lecture time on the last week of the class.
- We will be using the Honorlock system to monitor the quiz.
- Details of the quiz and the grading rubric will be posted on eLearning.

Course Outline/Academic Calendar

Week of	Topic/Lectures	Details/Materials	Assignment/Activity
Week 1 1/15/2024	Course access and self-introduction	Syllabus	Lab 0 – Python environment set-up
	Module 1: Introduction to Pandas, Python environment set-up	Scikit-learn: Machine Learning in Python	Lab 1
Week 2 1/22/2024	Module 2: Introduction to Machine Learning	H2O-3 – documentation	Lab 2
Week 3 1/29/2024	Module 3: Supervised Learning	• Supervised Learning – types and approaches • Decision trees	Lab 3
Week 4 2/5/2024	Module 3 continued	Linear models	**Project 1 – Posted** Lab 4
Week 5 2/12/2024	Module 3 continued	Intro to Ensemble Learning. Random Forest	Lab 5
Week 6 2/19/2024	Module 4: Model Tuning, Evaluation and Selection		Lab 6
Week 7 2/26/2024	Module 4 continued		
Week 8 10/11/2023	Module 5: Ensemble Learning	Combining Multiple Learners	
Week 9 3/4/2024	Module 5 continued		
Spring Break			
Week 10 3/18/2024	Module 6: ML Interpretability		**Project 2 - Posted**
Week 11 3/25/2024	Module 6: continued		Lab 7
Week 12 4/1/2024	Module 7 – model debugging		
Week 13 4/8/2024	Module 8: Dimensionality Reduction	Dimensionality Reduction: PCA and GLRM (H2O)	Lab 8
Week 14 4/15/2024	Module 9: Neural Networks	• Perceptron Learning • Weighted Networks - The perceptron • The Perceptron: A probabilistic model for information storage and organization in brain	Lab 9
Week 15 4/22/2024	Overview of LLMs	Overview of Large Language Models and how to use LLM for Classification/Regression	

Course Policies

Make-up exams: No makeup exam is allowed in this course.
Extra Credit: There is not any extra credit activity for this course.

Late work policy

- All submissions, including labs and projects, will have a grace period of 10 hours.
- Submissions after the grace period will not be accepted.
- The grace period is intended to help students who experience technical problems with eLearning.
- The grace period is not intended to give students extra time to work on assignments.
- Plan to submit your work early if you know you will not around the due date.
    - Submission links will be available at least 1 week in advance of the due date.
    - This gives you plenty of time to submit your work early, even if you have personal commitments, work load, or UTD work load that will make it difficult to complete the assignment on time.

Class Participation and Camera use

- Optional: Class participation during Project introduction. The sessions will be recorded and conducted via MS Teams. Exact date/time and links to join will be posted in the eLearning
- Students are encouraged to participate in the discussion on class MS Teams channel.

Comet Creed

This creed was voted on by the UT Dallas student body in 2014. It is a standard that Comets choose to live by and encourage others to do the same:

"*As a Comet, I pledge honesty, integrity, and service in all that I do.*"

UT Dallas Syllabus Policies and Procedures

The information contained in the following link constitutes the University's policies and procedures segment of the course syllabus. Please visit UT Dallas Syllabus Policies webpage for these policies.

The descriptions and timelines contained in this syllabus are subject to change at the discretion of the Professor.