AEB 6933: Special Topics: Applied Machine Learning
Spring 2026

Instructor:
Gulcan Onel
1091 McCarty Hall B
gulcan.onel@ufl.edu
352-294-7657
Instructor office hour: Tuesdays 1:00PM – 2:00PM in-person, or by appointment on Zoom (TBD).

Teaching Assistant:
Fernando Brito
f.britogonzalez@ufl.edu
TA office hour: Wednesdays 10:00AM – 11:00AM on Zoom (TBD)

Class Meetings:
Tuesdays: 3:00 PM – 4:55 PM, In-person, MCCA 2196
Thursdays: 4:05 PM – 4:55 PM, On Zoom (TBD)

Course Description:
This course provides an introduction to modern machine learning methods and their applications in applied economics, finance, and related fields. Students will explore the entire machine learning workflow, including data collection, model development, evaluation, and deployment. Emphasis will be placed on practical implementation, reproducibility, and the clear communication of results. Using the R programming language, students will develop and deploy machine learning models, focusing primarily on tabular data and "shallow" machine learning techniques. While the course will briefly address deep learning, the emphasis will be on its meaningful integration into applied economics and data science projects. Students will learn to apply their skills to common problems in economics and other fields.

Learning Objectives:
This course is designed to be hands-on, and students are expected to learn how to apply different machine learning techniques for analyzing different types of data. In order to achieve this objective, each discussed topic is accompanied by a lab session in which we examine how to use that technique on a data set. Specific objectives include

- **Understand Core Concepts:** Develop a comprehensive understanding of key principles in statistical machine learning, including supervised and unsupervised learning, model evaluation, regularization, and dimensionality reduction.
- **Apply Machine Learning Techniques:** Gain practical experience in applying machine learning methods in the R programming language to analyze and solve real-world problems (Students who prefer to use Python are allowed to do so, though my solutions will be in R).
- **Interpret Model Results:** Learn to critically evaluate machine learning models, understand their outputs, and use evaluation metrics to make informed decisions.
- **Think Critically and Ethically:** Develop critical thinking skills to assess trade-offs in model complexity and generalization while considering the ethical implications and potential biases of machine learning models.

Prerequisites:
Students are expected to have completed a graduate-level course in econometrics, statistics, or data science, as these foundational concepts will be built upon throughout the semester. While prior experience with R software is helpful, it is not required. If you are unsure whether this course is a good fit for you, feel free to contact me via email at gulcan.onel@ufl.edu .

Recommended Textbooks:

James, G., D. Witten, T. Hastie and R. Tibshirani (2021). An Introduction to Statistical Learning with Applications in R. Second Edition. Springer. A free PDF version of the book is available at https://www.statlearning.com/ (*This is the main book we will follow*).

Hastie, T., R. Tibshirani and J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer. The PDF version of the book can be downloaded from the publisher through UF Libraries: https://link.springer.com/book/10.1007/978-0-387-84858-7 (*This is for those who like a deeper mathematical treatment of the concepts covered in the course*).

Course Canvas Page:
The course has a dedicated Canvas page, which you can access through https://elearning.ufl.edu/. I will post course assignments, applications, and lecture notes here. You will also submit your homework assignments and class project through Canvas. Please ensure that you have announcement notifications enabled in Canvas.

COURSE GRADE AND ASSIGNMENTS
The overall course grade will be determined by

- Homework Assignments (50% total, lowest score dropped)
- Class project (30%)
- Class project presentation (10%)
- Participation and attendance (10%)

Letter grades will be assigned to the overall course grade as follows:

A = 93 and higher	C+ = 77 – 79
A- = 90 – 92	C = 73 – 76
B+ = 87 – 89	C- = 70 – 72
B = 83 – 86	D+ = 67 – 69
B- = 80 – 82	D = 63 – 66
	D- = 60 – 62

For more information on current UF policies for assigning grade points, see https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx

Homework Policy:
Collaboration on homework is encouraged; however, each student must submit their own original work. If you discuss an assignment with peers, please include their names in your submission. Directly copying code or written work from others is not permitted.

Late submissions and make-up assignments will not be offered, but the lowest homework score will be dropped. If you miss a deadline, you may submit the assignment by the start of the next class (usually in 24 hours), subject to a 20% penalty.

You are allowed to consult online resources (e.g., *StackOverflow*) for coding examples unless explicitly stated otherwise. Any reused code must be adapted to your specific task, and the source must be cited with a comment (*e.g., # Source: StackOverflow*) near where the code appears. Failure to cite reused code may result in it being treated as plagiarism.

No Exams:
There will be no exams in this class. Instead, students will prepare and present an applied class project toward the end of the semester.

Class project:
For the final project, students may choose any topic of interest that incorporates machine learning methods and may work individually or in pairs. Each student or group will submit a concise report detailing the project's objectives, data sources, methods used, and results, and will also present their findings to the class. Additional instructions and example project ideas will be provided later in the semester.

Attendance:
Attendance and participation account for 10% of the overall course grade, and students are strongly encouraged to attend class regularly. If you expect to be absent, please notify me in advance whenever possible.

CLASS POLICY ON THE USE OF GENERATIVE AI (GAI)

Generative AI tools like ChatGPT and GitHub Copilot can be helpful for learning, but they must support your learning, not replace it. This course focuses on building your skills and understanding, so the following guidelines apply:

✅ **Allowed Uses**

- **Reference Tool:** Use GAI as you would documentation or web search, such as asking how to create a chart in R or write a specific function. Always evaluate and understand the results.

- **Coding Help:** You can use GAI to assist with writing code, but you must understand how it works. Example: Let's say you have a dataset and want to check for missing values. You might ask a generative AI tool like ChatGPT the following: "*How can I check for missing values in a dataset in R?"* You can then adapt the response to your specific dataset.

❌ **Prohibited Uses**

- Do not use GAI to complete assignments for you, such as analyzing data or generating full responses.
- Do not use GAI to write narrative sections unless explicitly allowed.

Accountability

You are responsible for your work. If needed, I may ask you to explain your submissions to confirm your understanding. Failure to demonstrate understanding may result in grade penalties. For example, using the missing values example above, you must understand that The *is.na()* function checks for missing values in a dataset.

University Policies and Academic Resources:
Students seeking information on academic accommodations, university services, or official UF policies should refer to the UF syllabus policy page: https://syllabus.ufl.edu/syllabus-policy/uf-syllabus-policy-links/.

TENTATIVE COURSE OUTLINE:
(Coverage may change as circumstances warrant. Chapter numbers refer to the 2nd edition of Introduction to Statistical Learning (e.g., ISL Ch#). Additional readings will be provided through Canvas, as needed.)

Introduction

Overview of Machine Learning [ISL Ch2] (Regression Models, Model Selection and Bias-Variance Tradeoff, Classification, Introduction to R)

Linear Regression [ISL Ch3] (Simple Linear Regression, Hypothesis Testing and Confidence Intervals, Multiple Linear Regression, Extensions of the Linear Model, Linear Regression in R)

Classification [ISL Ch4] (Classification Problems, Logistic Regression, Multivariate Logistic Regression, Discriminant Analysis, Generalized Linear Models, Classification in R)

Resampling Methods [ISL Ch5] (Cross-validation, The Bootstrap, Resampling in R)

Linear Model Selection and Regularization [ISL Ch6] (Best-Subset Selection, Stepwise Selection, Validation and Cross-Validation, Shrinkage Methods, Ridge Regression, The Lasso, Tuning Parameter Selection, Dimension Reduction Methods, Principal Components Regression, Partial Least Squares, Model Selection in R)

Moving Beyond Linearity [ISL Ch7] (Polynomials, Step Functions, Splines, Generalized Additive Models, Local Regression, Nonlinear Functions in R)

Tree-Based Methods [ISL Ch8] (Tree-Based Methods, Classification Trees, Random Forests, Bayesian Additive Regression Trees, Tree-Based Methods in R)

Deep Learning [ISL Ch10] (Neural Networks, Convolutional Neural Networks, Time-Series Forecasting, Deep Learning in R)

Unsupervised Learning [ISL Ch12] (Principal Components, Higher Order Principal Components, k-means Clustering, Hierarchical Clustering, Unsupervised Learning in R)