

OPAN 6602 - Machine Learning I
Classes: October 29 - December 3

Semester/Year	Fall 2024/Mod 2
Professor	Dr. Tommy Jones
E-mail	twj22@georgetown.edu TA: Sheeba Moghal (sm3924@georgetown.edu)
Schedule	Section 1 (HOYA): Tuesdays, 6:00pm to 7:20pm ET Section 2 (SAXA): Tuesdays, 8:00pm to 9:20pm ET
Office Hours	Dr. Jones: Wednesdays, 4:00 to 5:00 PM ET or by appointment Sheeba Moghal (TA): Mondays, 4 to 5:30 PM ET Sheeba Moghal (TA): Saturdays, 1 to 2:30 PM ET
Course Webpage	Canvas

Live Sessions

October 29	Introduction to Machine Learning and Modeling Features
November 5	Multiple Linear Regression
November 12	Regression Diagnostics
November 19	Logistic Regression
November 26	Regularized Regression Models
December 3	Decision Trees and Random Forests

Course Description: This first ML course in a two-part series will introduce you to the various concepts and applications around machine learning. In this course, we will introduce supervised and unsupervised techniques in ML and how these are applied in various business scenarios. The course includes core machine learning concepts such as data engineering (feature and target engineering) and modeling processes, including validation and testing of results that are applicable to all ML concepts. Specific ML techniques covered include regression (multiple linear and logistic), regularized regression techniques, biased regression models, decision trees, boosting, bagging, and random forests. We will cover all of these approaches in the context of marketing, finance, and other important business decisions. At the end of this course, you should have a basic understanding of how these methods work and be able to apply them in real business situations. We will use RStudio to run ML models and draw insights in this course.

Course Objectives

- Familiarizing students with the basic concepts in machine learning, including the modeling process, feature engineering, and target engineering.
- Understand the differences between supervised and unsupervised learning.
- Understand and apply various models and algorithms for supervised learning.
- Diagnose and correct for pathologies in supervised learning models.
- Further develop students' working knowledge of the R programming language, focused on implementing pipelines for supervised learning.

Course Materials and Resources: There is no required textbook for this course, however, the first book below, Introduction to Statistical Learning (ISLR), is considered a standard in the field. We will reference the R version of this book throughout the course.

1. *An Introduction to Statistical Learning*, (Gareth M. James, Daniela Witten, Trevor Hastie, Robert Tibshirani), Springer (Book's [website link](#))
 - a. The R version ([LINK](#))
 - b. The Python version ([LINK](#))
2. *The Elements of Statistical Learning* (2nd Ed.), (Trevor Hastie, Robert Tibshirani, Jerome Friedman), Springer-Verlag ([LINK](#))
3. R for Data Science, (Hadley Wickham, Mine Çetinkaya-Rundel, Garrett Grolemund), O'Reilly ([LINK](#))

Software: You will be using the R programming language and the RStudio IDE in this course. Check the Course Overview page on the course Canvas page for installation notes.

Course format and preparation for class: You are required to review the class recordings and read assigned materials before every live session. Live class sessions will consist of concepts, and practice work using R & RStudio. To be “active learners” and derive the most benefit from class discussions, please make sure to follow all assigned materials before every class. During live class sessions, please contribute to discussions and participate in hands-on work to familiarize yourself with the techniques and tools we study.

Canvas Website: There is a website for the course in Canvas containing required course content, including videos and interactive media, and other resources. Announcements, handouts, and additional resources will be posted. Please visit the site frequently and check for new course announcements and postings. It is your responsibility to keep up with all announcements made on the course website and take the appropriate actions.

Course Schedule

Pre-Course Work: Refer to Canvas for a detailed list of topics and pre/post work for each session.

October 29	Live Session: Introduction to Machine Learning and Modeling Features
November 5	Multiple Linear Regression
November 12	Regression Diagnostics

November 18	Project 1 Due
November 19	Logistic Regression
November 26	Regularized Regression Models
December 2	Project 2 Due
December 3	Decision Trees and Random Forests
December 16	Project 3 Due

Grading Policy: The final grade will be on the following assignments (Check the last page for details):

- Project 1: Multiple Linear Regression (Individual) 30%
- Project 2: Logistic Regression (Individual) 30%
- Project 3: (Group) 35% (15% report, 10% team presentation, 10% team peer-evaluation)
- Class participation (Individual) 5%

Grades and Feedback on Assignments: Grades and feedback will be posted on Canvas. Participation will be graded based on concept review quizzes given during each in-person meeting. Quiz results will not be graded, however *participation* in the quiz will factor into each student's participation score.

Course Deliverables:

- **Individual Projects:** Main body and citations should be 8 pages or less. You may include up to 5 pages of appendices for charts, tables, etc. Report should be written in 12 point font with 1.5 line spacing
- **Group Project:**
 - **Report:** Main body and citations should be 8 pages or less. You may include up to 5 pages of appendices for charts, tables, etc. Report should be written in 12 point font with 1.5 line spacing
 - **Presentation:** A recorded presentation lasting not more than 10 minutes clearly stating the business/real-world problem, your data, modeling approach, results, and business/real-world conclusion. (I recommend using Zoom to record your presentation.)
 - **Peer Evaluation:** A single page clearly stating your contributions, your teammate's contributions, and constructive feedback for your teammate on areas for improvement.

Communication Expectations:

- **Communication with Peers:** You will be expected to engage with your peers via the discussion board and other required tools on a regular basis.
- **Communication with Professor:** Please feel free to email the instructor and the TAs with your questions, concerns, and/or to schedule a time to meet over Zoom. When sending emails please remember to follow the guidelines outlined below.

If you have general questions about assignments and course materials, please post these questions in the General Question Discussion Board Forum which you can access by clicking Discussions in the course navigation menu. This is an open forum, and you are encouraged to give answers and help each other.

- *Check the syllabus.* Before sending your email or message, be sure that your question has not already been addressed in the syllabus or announcements.
- *Be patient.* If you have a concern and send me a message, you can expect a response within 2 business days. Please allow 7 business days for assessment submission feedback.
- *Specify subject.* Subject line should include the topic of the message and class title.

Adherence to Policies: By enrolling, you agree to follow all course policies, which I will treat as a binding contract. Please review them carefully, as there will be no exceptions.

Grade Appeals: If you believe there's a grading error, submit a written request within one week of receiving your grade. After this period, all grades are final.

Attendance: Consistent attendance is strongly encouraged, as it correlates with better performance. Prepare for each class by completing assigned readings and analysis. If you miss class, inform me and catch up independently, including borrowing notes from classmates.

Seeking Assistance: If you're struggling, use available resources and attend office hours. If you can't make office hours, schedule an alternative meeting.

Classroom Etiquette & Honor Code: Professionalism is expected, including active participation with video on during live sessions. Be punctual, avoid distractions, and adhere to Georgetown's Honor Code (<http://honorcouncil.georgetown.edu/>). Violations, including cheating and plagiarism, will be reported. Assignments may be checked via SafeAssign for plagiarism.

The syllabus cannot identify all possible solutions which represent honor code violations and is not meant to be comprehensive in this regard. Students are expected to abide by the intent and the letter of the honor code and are required to report any honor code violations that are observed.

Group Assignments:

- Work collaboratively within your group; one submission per group.
- Group members receive the same grade, except peer evaluations.

Time Commitment: Plan to spend 8-12 hours per week on each online module, matching the rigor of face-to-face courses.

For full policies on individual and group work, refer to the MSBA Program page and the Student Policies Document. (<https://georgetown.instructure.com/courses/179942/modules>)

Max Mean Grade Policy: The course follows McDonough's graduate school policy of the maximum mean Grade Point Average (GPA) of 3.5.

Instructional Continuity: Please see the Instructional Continuity section of the MSBA Student Syllabus Policies document. Deadlines for cases and homework submissions will not be affected by the university closure, unless otherwise notified.

Academic Accommodations: Please see the Academic Accommodations section of the MSBA Student Syllabus Policies document.

Religious Observances: Please see the Religious Observances section of the MSBA Student Syllabus Policies document.

Copyright of Course Materials: Please see the Copyright of Course Materials section of the MSBA Student Syllabus Policies document.

Netiquette Guidelines: Respect others' opinions and express your views courteously. Meaningful discussions come from open-mindedness and understanding, not antagonism. Civility is key—everyone has valuable insights based on their experiences, and we can all learn from each other.

Office of the Student Ombuds (OSO): The OSO offers confidential, impartial support for students facing university-related issues. Whether you need help with academic concerns, interpersonal conflicts, policy clarification, or sensitive situations like harassment or bullying, the OSO is a safe resource.

To make an appointment, email studentombuds@georgetown.edu or call 202-784-1081. More information is available at <https://studentombuds.georgetown.edu>.