DNSC6314: Machine Learning I

School of Business
The George Washington University

ADMINISTRATIVE INFORMATION

Instructor:	Dr. Long He	Office:	Funger 415C
	Associate Professor	Office Hours:	By appointment
	Department of Decision Sciences	Email:	longhe@gwu.edu
	School of Business	Website:	https://business.gwu.edu/long-he
	George Washington University		
TA:	Mr. Dhairya Shah	Email:	d.shah@gwmail.gwu.edu

COURSE OUTLINE

This course aims to foster a comprehension of the fundamental principles of data science, with a specific emphasis on utilizing machine learning (ML) as a powerful tool for addressing business challenges. Our approach involves the practical utilization of analytics tools, such as scikit-learn in Python. While the primary focus is on the introduction of fundamental ML models, the objective is to establish a connection between theoretical classroom knowledge and practical business applications. Throughout the course, we will also discuss analytical modeling techniques crucial for informed business decision-making, highlighting the role of data in facilitating sound and effective business solutions.

No.	Topic	Text
1	Introduction to Data Science o Supervised vs Unsupervised Learning o Bias-Variance Trade-off	[ISLR §2]
2	Regression: Linear Models o Linear regression o Non-linear Transformations	[ISLR §3]
3	Classification: Linear Models o Logistic Regression o Support Vector Machine	[ISLR §4, §9]
4	Model Performance o Cross-validation	[ISLR §5]
5	Model Selection and Regularization o Subset Selection o Shrinkage Methods	[ISLR §6]

References:
– [ISLR] An Introduction to Statistical Learning --- with Applications in R [Second Edition]
   o   https://hastie.su.domains/ISLR2/ISLRv2_website.pdf

- [DSB] Data Science for Business
  - http://www.data-science-for-biz.com/

Software: Python Programming

Resources:
- ChatGPT Prompts for Data Science
- Youtube@StatQuest

BLACKBOARD (Q&A)

Active participation via the Blackboard course page is required. It is crucial to regularly check Blackboard for important updates, including announcements, handouts, the course schedule, classwork, homework assignments, exams, and other materials. The Blackboard course page also features a Discussion Board where you can engage with your classmates and communicate with me about course-related topics. To stay informed, I recommend subscribing to the discussion forums to receive notifications of new postings.

Please note that I prioritize responding to questions posed through Blackboard. Course-related questions sent via email will generally not be answered unless they are private and not relevant to the rest of the class.

ASSESSMENT

Final grades will be assigned based on the following components:

Component	Weight
Individual Assignments	70%
In-class Quiz (Feb 27, 2025)	30%
Total	**100%**

Grade Bands (These are approximate and subject to adjustments if necessary):
- **A**: 92.5% and above
- **A-**: 85% to < 92.5%
- **B+**: 80% to < 85%
- **B**: 75% to < 80%
- **B-**: 70% to < 75%
- **C+**: 65% to < 70%
- **C**: 60% to < 65%
- **C-**: 55% to < 60%
- **F**: Below 55%

INDIVIDUAL ASSIGNMENT

Assignments will include both conceptual and practical (e.g., coding). While you will complete the assignments individually, you are permitted to use GenAI tools (e.g., ChatGPT). Please refer to "GenAI Guidelines" below. The point values for each assignment may vary. For example, if there are three assignments, and you score 80 out of 100 on Assignment 1, 150 out of 150 on Assignment 2, and 100 out of 120 on Assignment 3, your total points for the Assignment Component of your final grade will be calculated as $(Assignment\ Weight) \times \frac{(80+150+100)}{(100+150+120)}$.

Late Submission: Deliverables must be submitted through Blackboard by the specified due date and time, unless stated otherwise. Extensions will only be granted for valid reasons (e.g., medical emergency, family emergency, religious observances, athletic competitions). Examples of invalid reasons include but not limited to travel plans, social events, computer malfunctions, and pet sitting. Unexcused late submissions will incur a penalty of 10% per day (calculated in whole days: 1 day late, 2 days late, etc.), including weekends and holidays, for up to three days. Submissions received more than three days after the due date will not be accepted and will receive a grade of zero.

QUIZ

There will be one in-class quiz. Please mark your calendar, as you are expected to take the quiz during class time. Makeup quizzes will only be allowed for valid reasons.

The quiz will be *open* book and consists of True/False questions, Multiple-Choice questions, and calculation problems. Calculators are permitted, but the use of computers, tablets, or phones is prohibited. If you do not have a calculator, you can still earn most points by providing correct formulations, as correct answers could be derived using a calculator.

GenAI Guidelines

You are encouraged to explore generative artificial intelligence (GenAI) tools, such as ChatGPT, to enhance your learning experience. At times, you may be explicitly asked to use ChatGPT and analyze its output for specific problems. When utilizing GenAI tools, it is essential to appropriately ***acknowledge*** and ***cite*** their use in your written submissions (e.g., code, reports), including the specific version of the tool employed. Please note that content generated by GenAI tools may not always be accurate or suitable. It is your responsibility to critically evaluate the validity and applicability of any GenAI output included in your submissions. Failure to adhere to these guidelines will be considered a violation of the GW Code of Academic Integrity. Please note that definitions and expectations regarding "plagiarism, cheating, and acceptable assistance" may differ across courses and instructors.

CODE OF ACADEMIC INTEGRITY

Academic integrity violations will be reported. For detailed information, please refer to
https://students.gwu.edu/code-academic-integrity.