BOSTON UNIVERSITY

Module 3 – Machine Learning Fundamentals

Week 1: Introduction and Welcome

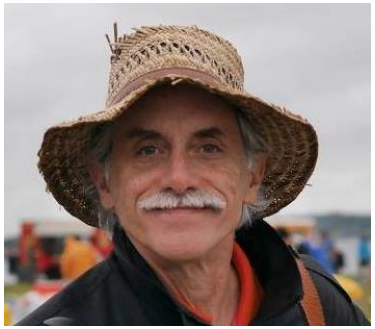
BU Faculty of Computing & Data Sciences

Agenda for Today

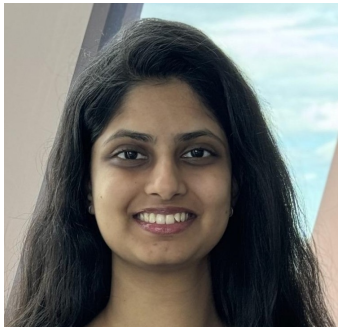
- Teaching Team
- Schedule
- Learning Objectives
- Topics
- Grading Policies
- Resources
- Q & A

Teaching Team

Instructor Wayne Snyder



Learning Facilitators



Parita Danole

David Kim

Will Mackin

Semester Schedule: Live Sessions, Homeworks, and Project

8 Weekly Homeworks

Project

Milestone 1

Milestone 2

Final Report

Week	Live Sessions Mon 7-8pm	HW Due Date Sun 11:59	Topic
Week 1	1/21 (Tues 8-9)	1/26	Introduction to Machine Learning
Week 2	1/27	2/2	Linear Regression
Week 3	2/3	2/9	Generalization, Errors, and Bias-Variance Tradeoff
Week 4	2/10	2/16	Training, Testing, and Validation Sets; Cross-Validation
Week 5	2/18 (Tues 8-9)	2/23	Feature Selection and Regularization
Week 6	2/24	3/2	Decision Trees
Week 7	3/3	3/9	Ensemble Methods: Bagging, Random Forests, Boosting Trees
Week 8	3/10	3/16	Classification
Spring Break			
Week 9	3/24		Variable Importance Measures
Week 10	3/31	4/6 (M 1)	Causal Inference
Week 11	4/7		Bias and Fairness
Week 12	4/14	4/20 (M 2)	Final Project Management
Week 13	4/22 (Tues 8-9)		Unsupervised Learning
Week 14	4/28	5/2 (Final)	Preview: Deep Learning

Learning Objectives

This is the first in a two course sequence in machine learning:

Module 3: Fundamental concepts and classic (non-deep) algorithms

Module 7: Deep Learning with Artificial Neural Networks

Our objectives this term:

- Understand the main **types of machine learning**: supervised vs unsupervised, regression vs classification.
- **Implement** basic machine learning **models**, including linear and polynomial regression, decision trees, and ensemble methods.
- Appreciate the importance of **data preprocessing**, including data cleaning, feature engineering, and scaling.
- Develop **workflows for training, evaluating, and tuning** machine learning models.
- Explain the concepts of **underfitting and overfitting**, and learn techniques to address these issues.
- Navigate **tradeoffs** between model performance, interpretability, and computational efficiency.
- Recognize, analyze, and mitigate **bias and fairness** issues in machine learning, emphasizing ethical considerations.
- Effectively **communicate** machine learning results to technical and non-technical audiences.

Weekly Topics

Introduction to Machine Learning
Linear Regression
Generalization, Errors, and Bias-Variance Tradeoff
Training, Testing, and Validation Sets; Cross-Validation
Feature Selection and Regularization
Decision Trees
Ensemble Methods: Bagging, Random Forests, Boosting Trees
Classification
Variable Importance Measures
Causal Inference
Bias and Fairness
Final Project Management
Unsupervised Learning
Preview: Deep Learning

Grading Policies

Category	Grading Breakdown
Academic Content	500 points = 100%
• 8 Weekly Assignments	390 = 78% (40 for HW01, 50 for rest)
• Final Project	100 pts = 20%
• Milestone One	25 pts = 5%
• Milestone Two	25 pts = 5%
• Final submission	50 pts = 10%
• Discussion (Yellowdig)	10 pts = 2%

Late Homework Policy (same as fall term):

- **Due Date:** Homeworks are due by **Sunday midnight (1 minute after 11:59 PM)**.
- **Grace Period:** You have a **2-hour grace period** until **2:00 AM on Monday** to submit without penalty.
- **Late Submissions:**
  - Submissions made after 2:00 AM on Monday will incur a **10% penalty per day late**.
  - **No homework will be accepted after 5 days** (i.e., midnight on Friday).

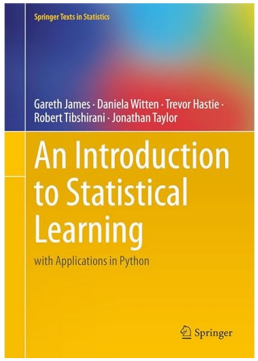
Homework Submission and Grading Guidelines

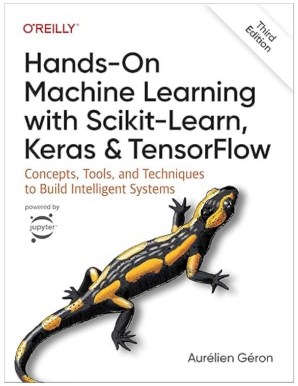
We will use **Gradescope** for homework submission, grading, and addressing any grading-related questions or concerns.

- **Resubmissions**:
  - You may resubmit your homework as many times as needed until the end of the late submission period. Please note that the **submission date and time will be based on your most recent upload.**
- **Grading**:
  - Homework assignments will be **autograded**.
  - For the first two homework assignments, autograding results will be available immediately after submission to help you acclimate to the system. For all subsequent assignments, grades will be released after the late submission period ends.
  - The three project submissions will be **manually graded** by the LFs.
- **Regrades**
  - Gradescope gives you the opportunity to communicate with us about potential problems in grading.
  - PLEASE only use this if you think there has been a mistake, not to argue about our grading policies.

Resources

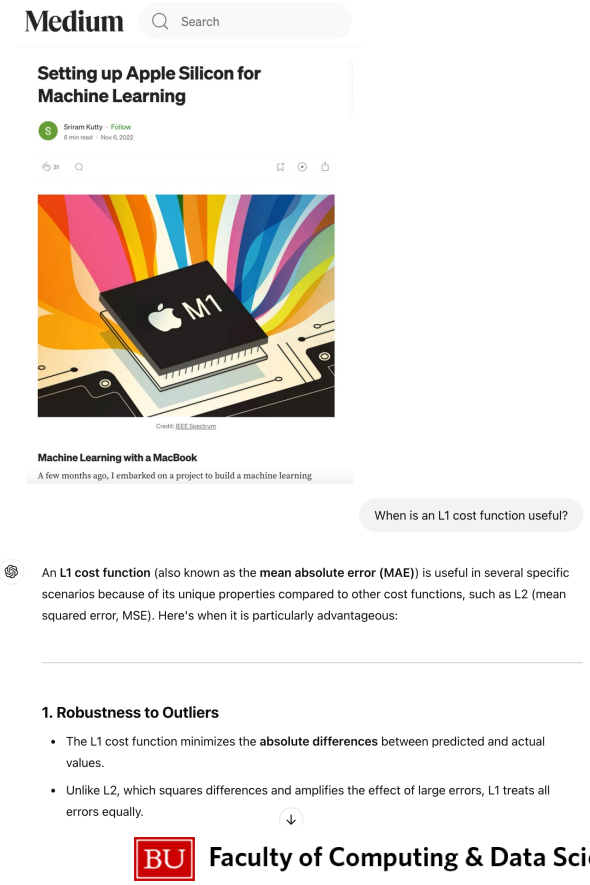
Textbooks (both are available online or through the BU library):





I encourage you, but do not require, to subscribe to **Medium** ($5/month) for a well-curated collection of short blog posts about machine learning. I'll post PDFs of those I refer to in the lessons.

I also recommend you sign up for a Plus account with **chatGPT** ($20/month), and use it to answer questions, find datasets and resources, tutor you, but not write your code (since you won't learn much if you do that).



Questions?

Do you have any questions?