CS167: Machine Learning

Syllabus
Logistics
Machine Learning: Introduction

Monday, August 24th, 2026

https://analytics.drake.edu/~reza/teaching/cs167_fall26/cs167_schedule.html

Introduction

https://analytics.drake.edu/~reza

Alimoor **Reza**

- Associate Professor of Computer Science, Dept. of Mathematics and Computer Science, Drake University
- Office: 323 Collier-Scripps
- Email: md.reza@drake.edu
- Phone: 515-271-1972
- Office hours: Tues/Thurs: 1:00-3:30pm CDT
  additionally by appointment
  Zoom link

Qingdao Drake United College, China Experience

Course overview

- Brush up on necessary programming tools and libraries
  - Python basics, Pandas library, PyTorch, GitHub, Google Colab (notebook), Kaggle (notebook)

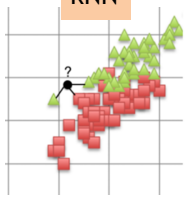
- Classical Machine Learning
  - K-nearest-neighbors (k-NN)
  - decision trees
  - random forests (RF)
  - perceptrons
  - support vector machines (SVM)

Generic Classical ML Pipeline

Training Set

Testing Set

kNN



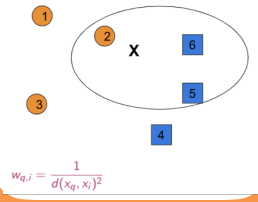
Evaluation Metrics

- Classification metrics
- Regression metrics

Training Set

Testing Set

Weighted kNN



Evaluation Metrics

- Classification metrics
- Regression metrics

Training Set

Testing Set

Decision Tree

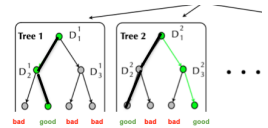
Evaluation Metrics

- Classification metrics
- Regression metrics

Training Set

Testing Set

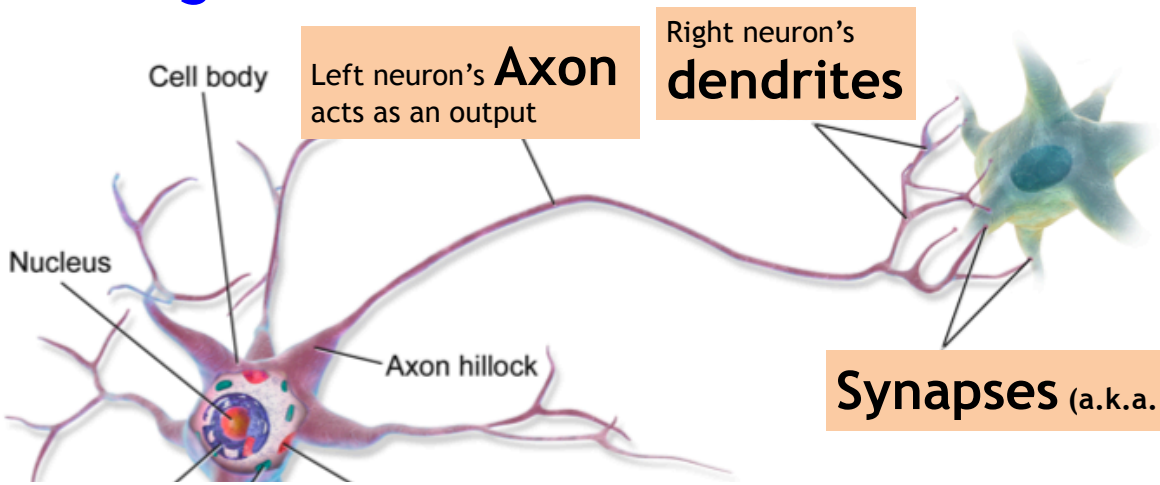
Random Forest



Evaluation Metrics

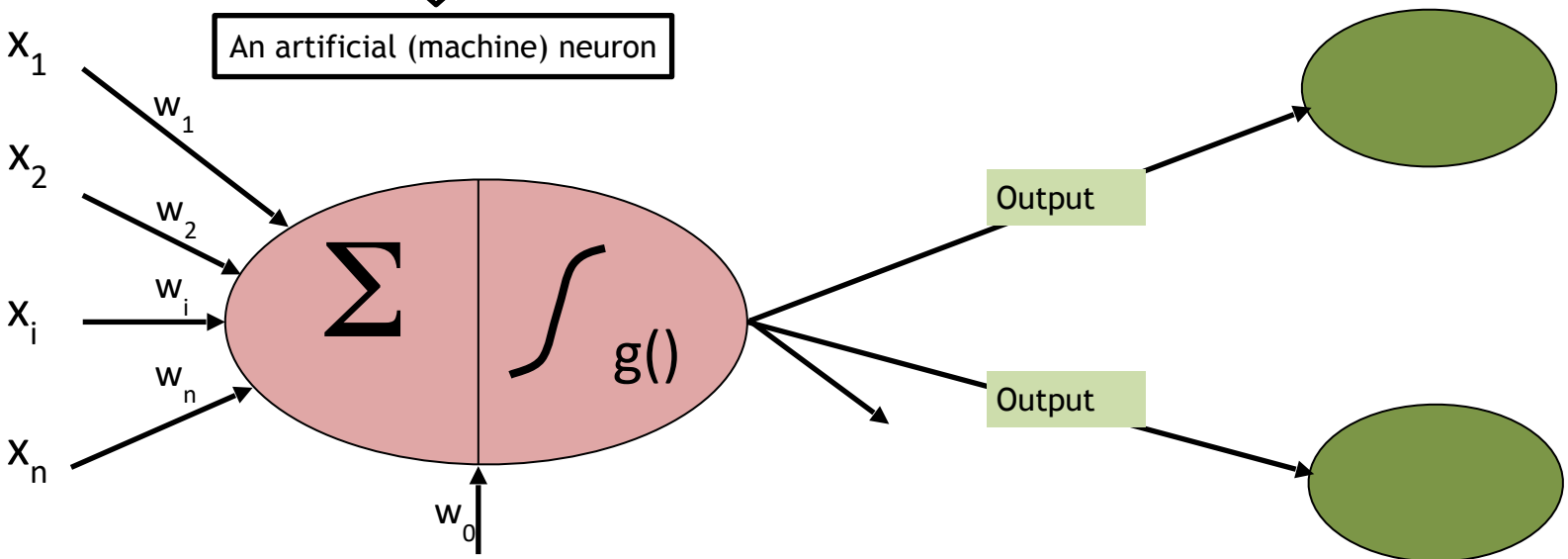
- Classification metrics
- Regression metrics

Deep Learning: Natural neurons vs. artificial neurons



Natural neuron

An artificial (machine) neuron



Course overview

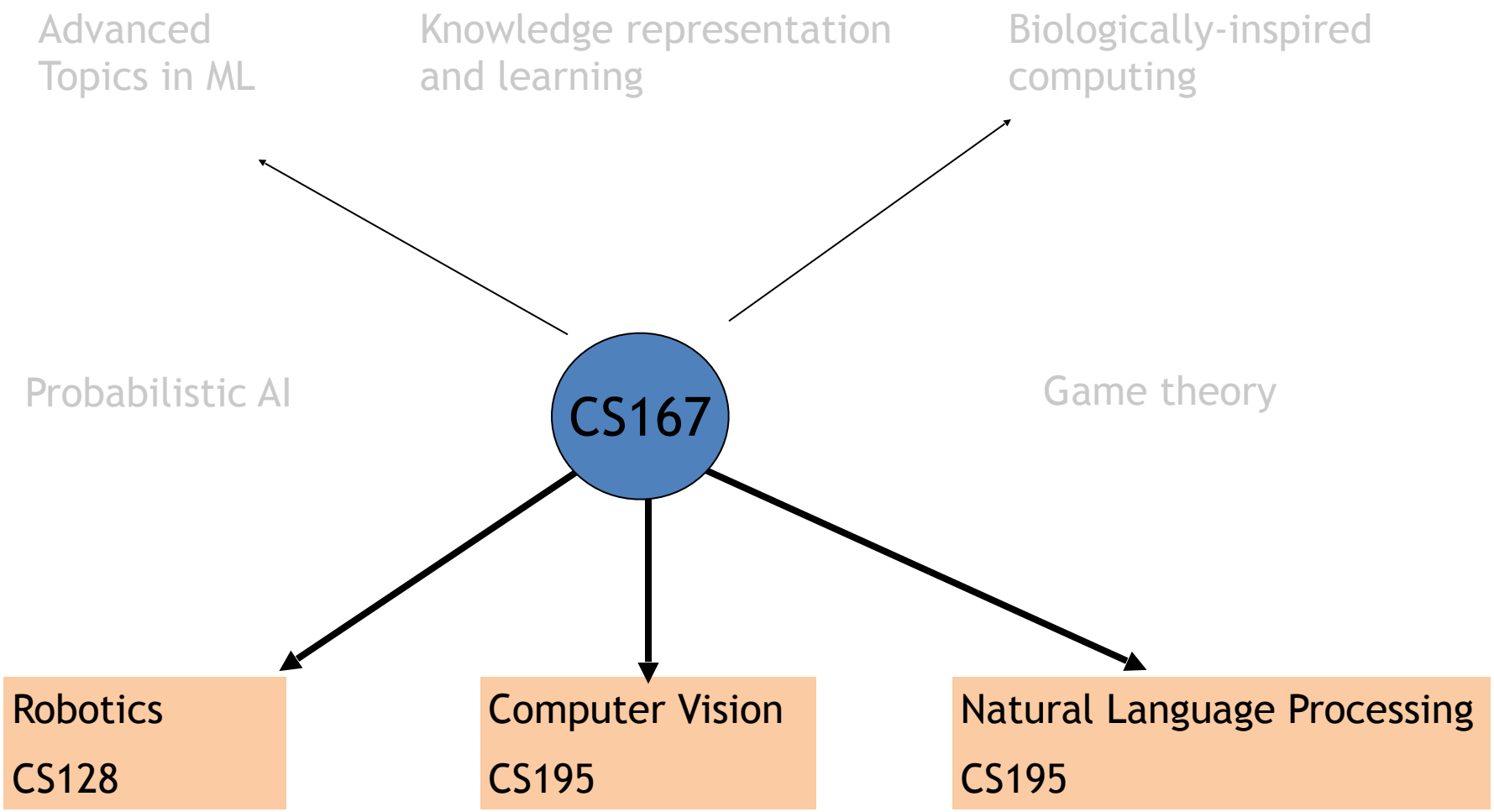
- Deep Machine Learning
  - Multilayer Perceptron (MLP)
  - Convolutional Neural Network (CNN)
  - Recurrent Neural Network (RNN)
  - Long Short-Term Memory (LSTM)
  - Transformers
- Applications
  - Computer vision, natural language processing, robotics, audio/speech analysis

Why take this class?

- Learn a lot about Machine Learning

  - Apply a variety of modeling techniques to classification, regression, and unsupervised learning problems using data in different formats (such as typical structured data, text, and images).
  - Create software that utilizes machine-learning programming libraries in order to conduct machine-learning-based data analysis.
  - Develop and conduct machine-learning-based data analysis experiments, and they will be able to interpret and explain the results.
  - Feel comfortable with using industry-standard tools such as Google Colab, GitHub, etc
  - Understand fundamentals of machine learning
  - Gain an understanding of the advantages and disadvantages of different learning paradigms so that students can choose appropriate solutions given a problem description
  - Receive hands-on experience with commonly used algorithms and software tools within machine learning

Why take this class?



Careers in ML/AI

- 'Pure' ML/AI
  - Academia, industry labs
- Applied ML/AI
  - Almost any area of CS
  - NLP, vision, robotics
  - Economics
- Cognitive Science

NAVAL RESEARCH LABORATORY
WASHINGTON, DC

LOCKHEED MARTIN

iRobot®

SRI International

Microsoft Research

Google™

Prerequisites

- CS 66: Introduction to Computer Science II
- Practically, this means:
  - Proficiency in a general-purpose programming language
  - Some level of mathematical maturity will be helpful, esp. with calculus, linear algebra, statistics
  - Willingness to learn some programming and/or math on your own if necessary

Course Logistics

Contacting

- I am here to help! To get a reply as quickly as possible:
  - For questions about the class, assignments, personal matters, etc: email md.reza@drake.edu

Course mechanics

- Syllabus, schedule, assignments, announcements, etc. on Drake Blackboard
  - https://drake.blackboard.com
  - Syllabus link

- Class meeting times and locations:

  - Time: Monday/Wednesday 2:00 pm - 3:15 pm (CRN# 8363) and Location: Collier-Scripps # 301
  - Time: Monday/Wednesday 3:30 pm - 4:45 pm (CRN# 12639) and Location: Collier-Scripps # 335

Course mechanics

- Office hour times and location:

  - Office hour#1: Tuesday: 12:00 pm - 2:30 pm
  - Office hour#2: Thursday 12:00 pm - 2:30 pm
  - Location: Collier-Scripps#323

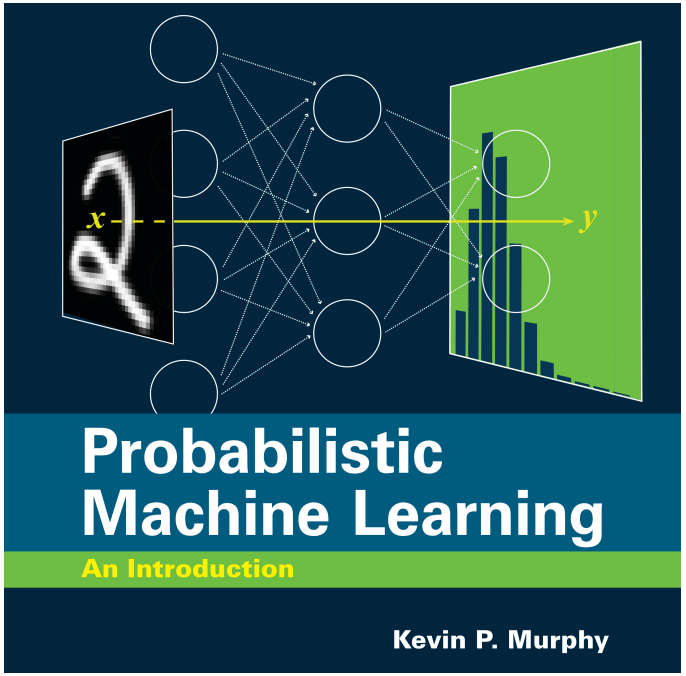
Additionally by appointment: Zoom link

Course mechanics

- Textbook:
  - We will not follow any particular textbook
  - Optional chapter reading will be assigned from:

    Probabilistic Machine Learning: An Introduction

    - Kevin P. Murphy (2022)



Grading and Requirement

- Notebook assignments (25%): *take home 5 assignments (submit on Blackboard)*
- In-class activities (20%): *10-12 simple coding activities or paper-based tasks (submit on Blackboard)*
- Midterm and Final Exams (30%):
- Projects (20%): *2 projects, each worth 10% of final grade*
- Attendance/Participation (05%): *Participation in polls, not based on correctness, physical attendance during lecture time*

Grading and Requirement

- Notebook assignments (25%): *take home assignments*.

Notebooks Assignments (25%) Throughout the course, we will be learning to apply machine learning principles using Python machine learning tools. Machine learning code is often developed in and communicated using an interactive integrated development environment called Jupyter Notebooks which support a natural interleaving of code, output/results, and mark-up documentation. You will regularly submit notebook files (files with the extension .ipynb) to demonstrate your proficiency with the Python tools we are using. Given the long computation times of the programs you write, I will not usually be executing your code, so it is critical that the results from your executions are preserved in the notebook. You can expect to submit **5 assignments** throughout the course.

Grading and Requirement

- In-class activities (20%): *8-9 simple coding activities or paper-based tasks (submit on Blackboard).*

In-class Activities (20%) You'll also regularly submit notebook files with simple coding or paper-based activities. These should be easier since they follow the lecture and are usually done during or right after class. I won't run your code, so please make sure to save the results in the notebook. Expect to submit about **8-9 in-class activities** throughout the course.

Grading and Requirement

- Midterm and Final Exams (30%):

Midterm and Final Exams (30%) The midterm exam will be administered via Blackboard during scheduled class time, while the final exam will also be administered via Blackboard but will take place during final exam week. Both exams will consist of the following question types: *(i) Multiple Choice Questions (MCQ), (ii) True/False, and (iii) Fill-in-the-Blank, some of which may require calculations.* Both exams are to be completed individually. Consistent with real-world practice, students may consult external resources, such as class notes and internet materials, during the exams. Any sources used, aside from class notes, must be properly cited.

Grading and Requirement

- Projects (20%): *2 projects, each worth 10% of final grade*

Projects (20%): The **2 projects** in this course will emphasize the design, execution, and interpretation of machine learning experiments. The grading emphasis will be on how well you explain your data and experiment as well as your written interpretation. For these, you will submit Jupyter Notebooks with more extensive writing in the mark-up cells than for your regular notebook assignments.

Grading and Requirement

- Attendance/Participation (05%): *Participation in papers + polls, not based on correctness*

Attendance/Participation (05%): This class is highly interactive, meaning that active participation is both expected and the norm. You will receive credit for your participation, and it will be counted towards your final grade. I will keep track of your involvement using a signature sheet. Throughout the course, I will pose questions using polling software and conduct in-class Q&A sessions to better understand how the class is grasping the content. These responses will not be evaluated for correctness but rather for completion. ***More than five absences may result in being withdrawn from the course.*** If you miss an assessment because of illness, you must provide documentation before you can take an assessment. All other absences from assessments must receive prior approval. Additionally, each late arrival (tardiness) or early departure will count as a half of an absence. Accumulating late arrivals or early departures will have the same impact as multiple absences on a student's final grade. It is each student's responsibility to keep track of their absences, late arrivals, and early departures to determine how it will affect their grade.

Grading Scale

- The tentative grading scale for this course would be as follows:

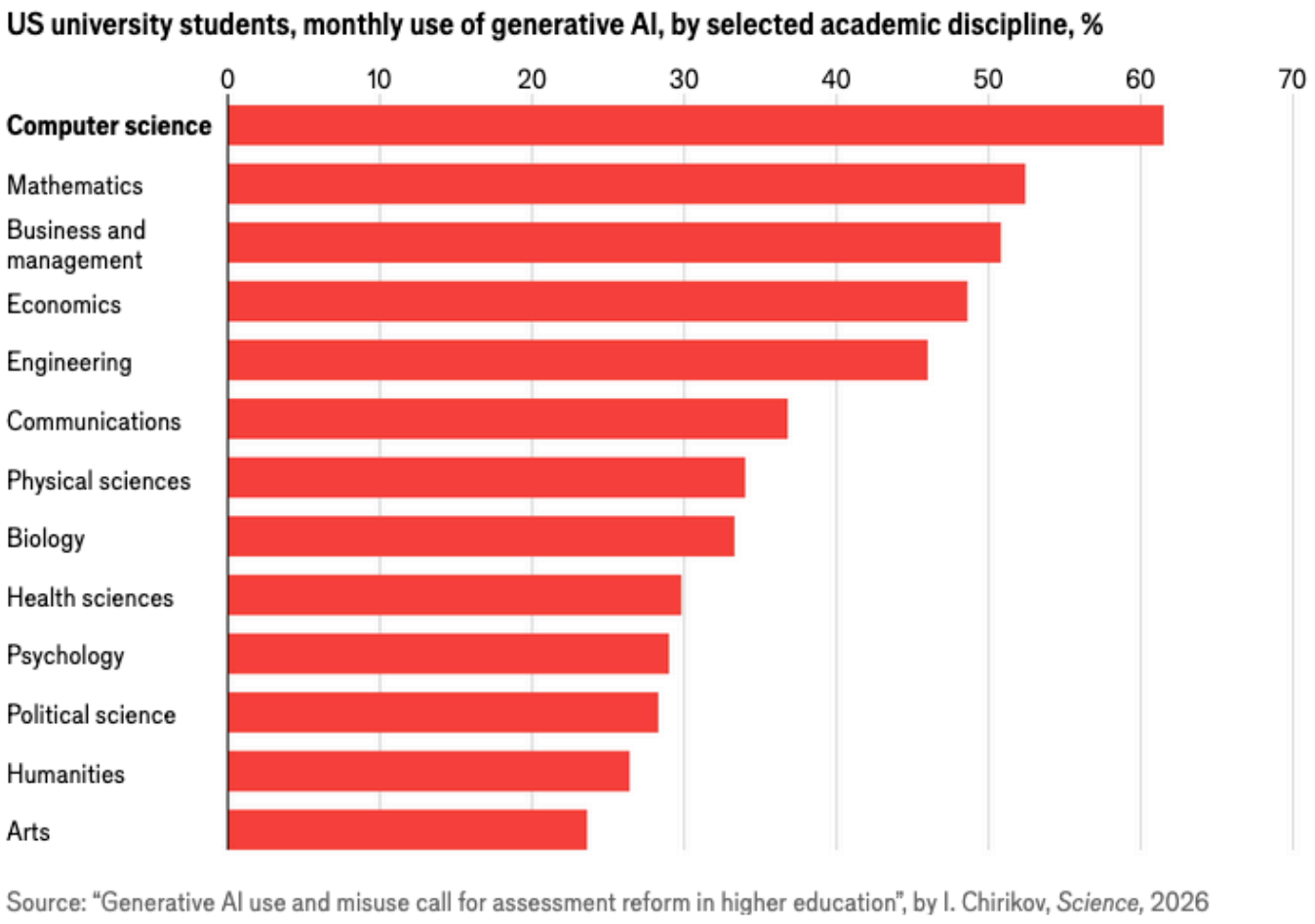
A (93%-100%)	A- (90%-92.9%)	B+ (87%-89.9%)
B (84%-86.9%)	B- (80%-83.9%)	C+ (77%-79.9%)
C (74%-76.9%)	C- (70%-73.9%)	D (60%-69.9%)
F (0%-59.9%)		

Grading and Requirement

- "Artificial intelligence may be accelerating these trends. Educators across every academic field are struggling to contain AI-assisted cheating. And computer-science students appear to be the heaviest users of such tools. A recent paper by Igor Chirikov of the University of California, Berkeley, and colleagues analyses survey data on generative-AI use among more than 95,000 undergraduates at 20 institutions. ***The researchers found that 62% of computer-science students use AI regularly, compared with 37% of students overall; one in ten AI-using computer-science students admitted to cheating with the technology. Online instruction makes such behavior harder to police***. Combined with all-you-can-eat course loads, AI may make it easier for students to race through programmers. It may not be a coincidence that computing degrees surged after Chat-GPT's release in 2022."

Reference: https://www.economist.com/graphic-detail/2026/06/01/do-you-really-want-that-computer-science-degree

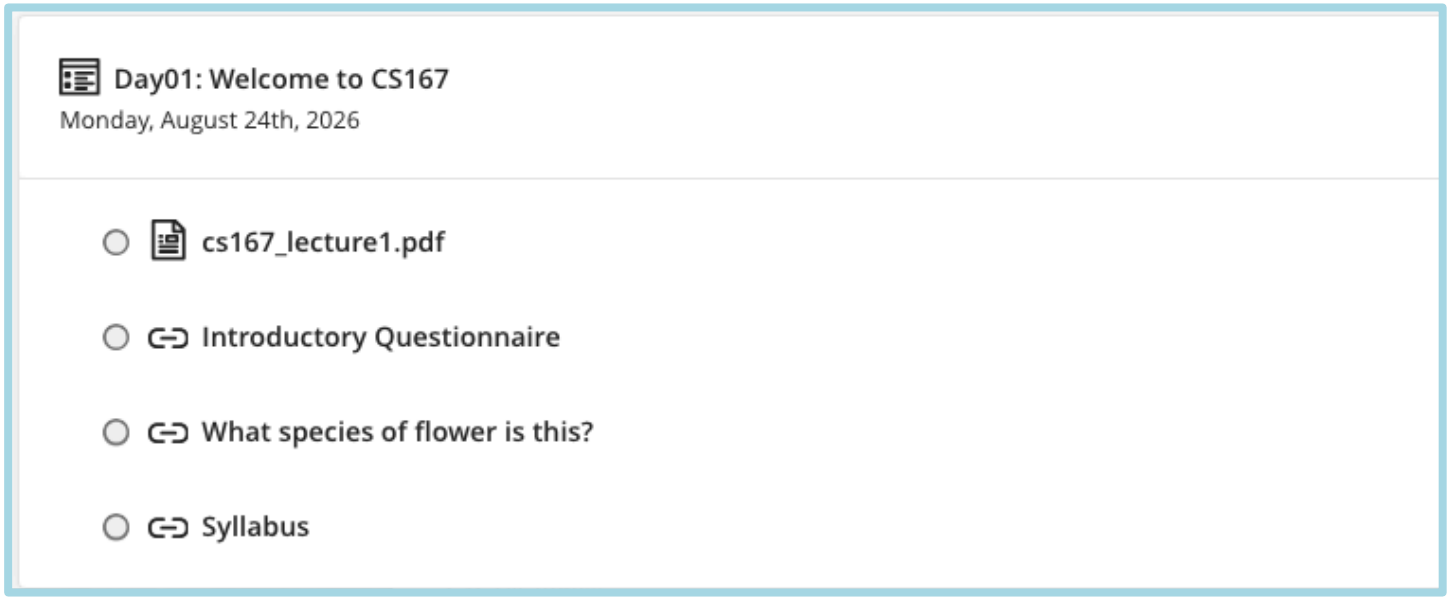
Grading and Requirement



Reference: https://www.economist.com/graphic-detail/2026/06/01/do-you-really-want-that-computer-science-degree

Blackboard Tour

Task: Posted on Blackboard



Introduction to Machine Learning

An Example: Classifying a Dog

- Imagine we want to classify which image depicts a specific dog we want to identify
- Our training data might look something like this:

An Example: Classifying a Dog

- Then, when we have some new pictures of one of the dogs, the idea is that we can make a **prediction** based on previous data as to whether it is **Zoey** or **Ace** in the photo.

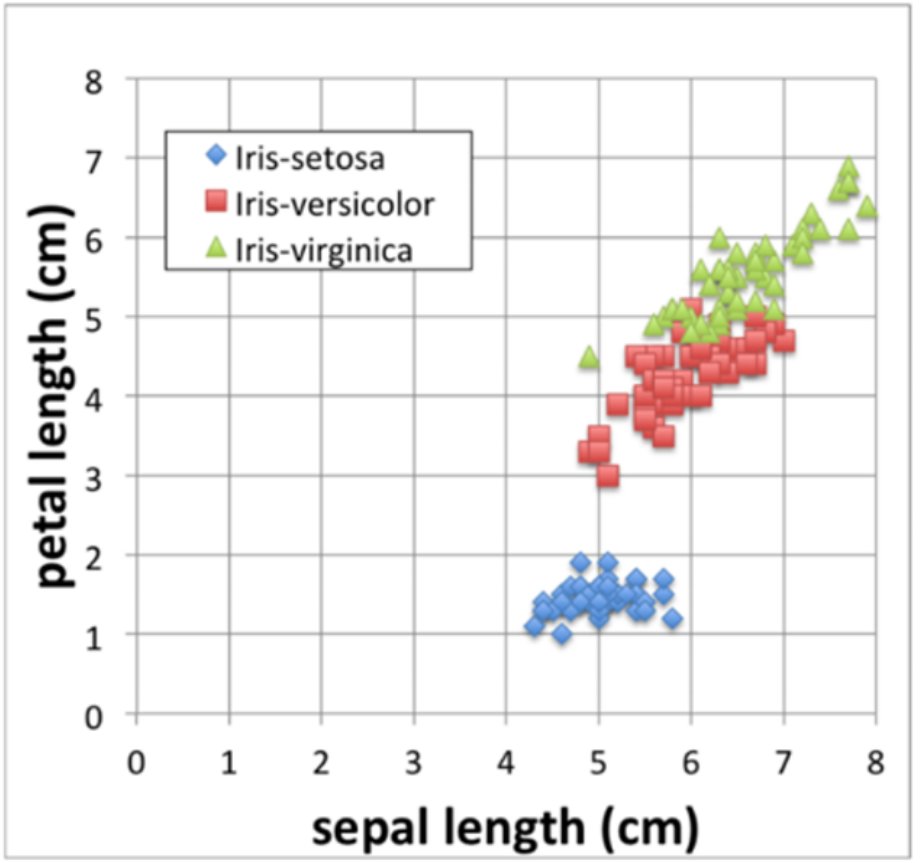
Another Example: What species of Iris is this?

- Use the poll link to answer this question

Imagine you found this beautiful flower while on a walk and took the following measurements:

5.1 cm petal length
7.2 cm sepal length

What species do you think it is?



Machine Learning Variations

- We are going to learn about a lot of different types of machine learning in CS167. Here are a few categories to look out for:

  - **classification:** identify which category it goes in, eg, *'Spam or ham?', 'Eric or Tim?', 'Fish, amphibian, reptile, bird, or mammal'*
  - **regression**: real-valued labels eg, price of Bitcoin, tomorrow's temperature, etc.
  - **supervised learning:** data has labels, goal is to predict the labels of new instance
  - **unsupervised learning:** data does not have a label, the goal is to analyze/ cluster the examples
  - **other issues:** missing data, sequential data, outlier anomaly detection, and many more

Group Exercise

- Group Exercise:
  - Take 2 minutes to talk to your neighbors about the following:
    - Come up with as many examples as you can of ways you interact with machine learning on a day-to-day basis.
    - Submit your answers to the following Google form:
      - https://forms.gle/w19WLdLkmEpnQ1HE9

In-class activity#0

- Blackboard Task
  - Try to finish **In-class activity#0** before next lecture but definitely by Friday

In-class activity#0

- Blackboard Task
  - Finish In-class activity#0 before Friday