ME 297: Introduction to Data Science for Mechanical Engineers

Instructor

- Ilias Bilionis, *Associate Professor of Mechanical Engineering, Purdue University*, ibilion@purdue.edu

Place and time

- Wednesday, 2:30-3:20 pm, at ME 1006

Office hours

- TBD Online

Teaching assistant

- Nimish Awalgaonkar, nawalgao@purdue.edu

Audience

Mechanical engineering sophomores

Course Description

This course introduces data science to sophomore mechanical engineers with no prior knowledge on the topic. We start by defining data science and demonstrating why mechanical engineers should care. Most data science nowadays happens in Python. So, the course starts with a brief overview of Python basics going from simple mathematical expressions to data loading and visualization. Data science must deal with uncertainty and randomness, and that is why we introduce several concepts from probability theory. We use these probabilistic concepts to summarize and compare datasets. Data science is about making models. For example, how can you make a model that will give you the efficiency of an engine given the operating conditions using experimental data? So, we end by showing how to use data to make models, and how to test if these models are good.

Course Learning Outcomes

After completing this course, you will be able to:

- Program in Python within a Jupyter notebook environment.
- Summarize and compare datasets using empirically estimated statistics.
- Summarize and compare datasets visually.
- Represent uncertainty using probabilities.
- Apply simple probability rules to propagate uncertainty.
- Estimate probabilities from data.
- Solve regression problems (learn from data a linear model that takes you from a set of input variables to a continuous variables).

- Solve classification problems (learn from data a model that takes you from a set of input variables to a discrete label).

Prerequisites

- Basic calculus.
- Matrix-vector multiplication.
- Some programming experience (e.g., Matlab or Python). You will learn Python as part of this course.

Required Textbook

There is a required textbook for this class. You can access it online for free here:

Ilias Bilionis, Introduction to Data Science for Mechanical Engineers (Lecture Book), 2021

The book is organized in 16 lectures. Each lecture contains the required theory along with programming Python sections that demonstrate the theory. You can (and should) run the code of the book on your own. For that you can either use Google Colab or Binder. Google Colab is probably easier to use, but it requires a Google account. To launch the code of a page, you click on the rocket button at the right top of a page and select the either Google Colab or Biner. It is also possible to run the code on your own computer, but I do not recommend as you will have to already be familiar with Python to do that.

Homework

This course relies heavily on homework. There is one required homework set due each week, before the lecture time. 100% of your grade comes from the homework. There are no quizzes or exams.
You can find the homework at the end of each lecture in our textbook. These are essentially Jupyter notebooks that you are supposed to launch in Google Colab and then submit them for grading. You will submit the homework through Brightspace. You will find specific instructions on how to do this in Brightspace.

Grading

Your final grade will be your average homework grade after dropping the lowest graded homework. The final grade will be turned into a percentage grade, and your letter grade will be based on the following scheme: A+ 98 – 100; A 88 – 98; A- 85 – 88; B+ 80 – 85; B 73 – 80; B- 70 – 73; C+ 67 – 70; C 62 – 67; C- 60 – 62; D+ 57 – 60; D 52 – 57; D- 50 – 52; F 0 – 50. The instructor reserves the right to change with grading scale.

Course Schedule

Period	Date	Topic	Reading	Homework
1	Aug. 25	Introduction to data science, Python basics	Lecture 1	Homework 1
2	Sep. 1	Data arrays, numerical Python	Lecture 2	Homework 2
3	Sep. 8	Data loading and selection	Lecture 3	Homework 3
4	Sep. 15	Data visualization	Lecture 4	Homework 4
5	Sep. 22	Functions, data manipulation, and models	Lecture 5	Homework 5
6	Sep. 29	Conditionals and loops	Lecture 6	Homework 6
7	Oct. 6	Probability as a measure of uncertainty	Lecture 7	Homework 7
8	Oct. 13	The basic rules of probability	Lecture 8	Homework 8
9	Oct. 20	Discrete random variables	Lecture 9	Homework 9
10	Oct. 27	Continuous random variables	Lecture 10	Homework 10
11	Nov. 3	Expectation and variance	Lecture 11	Homework 11
12	Nov. 10	Normal distribution, quantiles, credible intervals	Lecture 12	Homework 12
13	Nov. 17	Fitting models with the principle of maximum likelihood	Lecture 13	Homework 13
14	Nov. 24	Covariance, correlation, and linear regression	Lecture 14	Homework 14
15	Dec. 1	Linear regression continued	Lecture 15	Homework 15
16	Dec. 8	Classification via logistic regression	Lecture 16	Homework 16

Course Help

To get help with course content, comment in the Piazza discussion forums. Each homework assignment and lecture has its own corresponding discussion in the Piazza course site. By commenting in these discussion forums, the course team will be able to respond to your question more quickly.

Discussion Guidelines

Please follow the Discussion Guidelines when contributing to discussions in this course. Here are a few of the key points you should remember:

- Do not use offensive language. Present ideas appropriately.
- Be cautious in using Internet language. For example, do not capitalize all letters since this suggests shouting.
- Avoid using vernacular or slang language. This could possibly lead to misinterpretation.
- Do not hesitate to ask for feedback.
- Be concise and to the point.
- Think and edit before you push the “Send” button.

Accessibility Information

Purdue University strives to make learning experiences as accessible as possible. If you anticipate or experience physical or academic barriers based on disability, you are welcome to let me know so that we can discuss options. You are also encouraged to contact the Disability Resource Center at: drc@purdue.edu or by phone: 765-494-1247.

The **Office of Institutional Equity**, which is responsible for ensuring Americans with Disability Act compliance, can be contacted with any accessibility concerns at:

Phone: (765) 494-7253
Email: equity@purdue.edu
TTY: (765) 496-1343
Website

- Purdue's Disability Resource Center Website
- Purdue's Web Accessibility Policy
- Purdue Equal Access Frequently Asked Questions (FAQs)
- Google Accessibility
- Jupyter Notebook a11y toolbar (from Microsoft research)

Academic Integrity

Academic integrity is one of the highest values that Purdue University holds. Individuals are encouraged to alert university officials to potential breaches of this value by either emailing or by calling 765-494-8778. While information may be submitted anonymously, the more information that is submitted provides the greatest opportunity for the university to investigate the concern.

The Purdue Honor Pledge

"As a boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do. Accountable together - we are Purdue"

Nondiscrimination Statement

Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life. Link to Purdue's nondiscrimination policy statement.