

# AI 539: MACHINE LEARNING CHALLENGES IN THE REAL WORLD

Dr. Kiri Wagstaff | Tuesdays/Thursdays 12:00-1:20 p.m., KEC 1001 | 3 credits

## Instructor and TA

### Dr. Kiri Wagstaff

Office hrs:

**Tues./Thurs.:**

**1:30-2:30 p.m.**

(KEC 3043)

or by appointment

[kiri.wagstaff@oregonstate.edu](mailto:kiri.wagstaff@oregonstate.edu)

Web: <http://www.wkiri.com/>



### TA: Grace Diehl

Office hrs: **Friday 2:30-3:30 p.m.**

(KEC Atrium)

Email: [diehlg@oregonstate.edu](mailto:diehlg@oregonstate.edu)

## Logistics

No required textbook.

Readings will be provided.

You will need a computer to complete assignments.

### Canvas (course materials):

<https://canvas.oregonstate.edu/courses/1899467>

### Discord (discussion/Q&A):

<https://discord.com/channels/921157105585651742>

## Grading

10%: Warm-up reading

10%: In-class activities

30%: Try-it-out assignments

50%: Hands-on project

+ Opportunities for extra credit

A- 90-92; A 93 or greater

B- 80-82; B 83-86; B+ 87-89

C- 70-72; C 73-76; C+ 77-79

D- 60-62; D 63-66; D+ 67-69

F less than 60

(Grades are rounded)

## Why take this course?

We will explore the challenges that machine learning (ML) systems face when they move from the laboratory into the real world. We will be inspired by machine learning applied to problems from astronomy, planetary science, autonomous driving, criminal justice, marketing, etc..!

**Ask yourself:** What obstacles make machine learning hard to deploy?

## In this course:

We will identify ML challenges and develop strategies for addressing them. Specifically, you will:

1. Create a **data set profile** to provide early warning of possible issues
2. Employ techniques for handling missing data, class imbalance, sampling bias, noisy data, and other **data challenges**
3. Select appropriate **evaluation** methods given the data set properties and deployment (end-user) goals
4. Explore methods for model explainability and calibration to increase **user trust**
5. Employ methods for monitoring model performance, handling domain shift, and updating models to ensure **reliable deployment**
6. Apply and compare these methods on a **data set of your choice** and share your findings with the course community



## Prerequisites

This course is designed for graduate students with prior ML experience (CS 434 or CS/AI 534 or equivalent).

Required skills:

- Train and evaluate supervised learning models (e.g., using scikit-learn in Python or equivalent)
- Good programming skills, including design and documentation
- Good communication skills (for in-class discussions, assignment write-ups, and project presentation)



xkcd (Randall Munroe) <https://xkcd.com/1838/>

## Need Help?

- Ask a question on Discord
- If you send email, use a meaningful subject line that starts with "[AI539]". Email may take up to 24 hours for a response.



## Course Topics (Tentative)

- Data set profiling
- Missing values
- Class imbalance and sampling bias
- Meaningful evaluation
- Working with non-iid data
- Noisy data and labels
- Domain shift
- Deployment and maintenance
- Explainability
- Classifier calibration
- Active learning
- Adversaries
- Exploration and discovery

## Classmate Contact Info

1)

2)

3)

Dr. Kiri Wagstaff

## Expectations

### Of students:

- **Take charge of your success** (use resources, ask for help)
- Arrive prepared for and maintain focus during class meetings
- **Read each assignment's grading rubric** before submitting work and **ask questions** if the expectations are unclear
- Complete assignments on time
- Contribute to a **positive learning environment**
- Follow OSU Code of Conduct: <https://beav.es/codeofconduct>



### Of the instructor:

- Provide informative class experiences and opportunities for practice (assignments, in-class activities)
- Provide timely feedback on student work
- Respect student time & workload (**Note: 1 credit = 2 hours outside class**)
- Answer questions, provide guidance, and communicate clearly

## Coursework

- **Readings:** You will provide short responses to relevant readings that highlight machine learning challenges (and solutions).
- **Assignments:** Short assignments in the first half of the term will enable you to compare alternative solutions to a particular challenge.
- **Final Project:** This course is centered on giving you the opportunity to detect, characterize, and address challenges in using machine learning on a data set of your choice. You will choose a data set, identify its challenges, evaluate solutions, and share your results in a written report and brief in-class presentation.



## Academic Integrity

All work that you submit should be **yours alone**, except when explicitly indicated in the instructions. **You may not share your solutions, write-ups, or code; nor may you look at others' work.** If others ask to see your work, direct them to come to me or Grace (GTA) for help. For the final project, discussing your project and strategizing with others is **encouraged**, as long as **the work you submit for grading is your own.**

If you employ code found on the Web, **you must give credit** to the original author and identify which parts are yours versus theirs.

Ensure you are familiar with what is and is not allowed and the OSU Academic Integrity Process (see <https://studentlife.oregonstate.edu/studentconduct/academicmisconduct>).

**If you are stuck, ask me for help. I want you to succeed!**

[kiri.wagstaff@oregonstate.edu](mailto:kiri.wagstaff@oregonstate.edu)

## Late Work and Regrading Policies

To encourage planning ahead, **late assignments or project components incur a penalty of 2% for each hour they are late.** E.g., if you submit 5 hours late, the maximum score is reduced by 10%. **No late work is accepted after 25 hours (50% penalty) have elapsed.** If an issue such as an illness or life crisis will prevent you from submitting your work on time, email me **at least 8 hours before the deadline** to request an extension. I will do my best to work with you.

If you spot a grading error, email me **within 48 hours** of the grade posting on Canvas so I can review it. The entire assignment will be regraded, not just individual parts, so the total score could go up or down.

## COVID-19 Safety Precautions

**Vaccines** are required by OSU policy. They have been shown to effectively reduce viral transmission. Exceptions must be approved through Disability Access Services (let me know if you need guidance). Masks, boosters, and self-tests are available. More info: <https://covid.oregonstate.edu/>

If you **test positive**, isolate for 5 days. If you don't feel well, please stay home and rest. Contact me to discuss catching up. Use available resources (office hours, Zoom, GTA, Discord, ask classmates for notes).

Please give everyone **more personal space** than you're used to from pre-COVID times.

**Let me know if you need a mask or any other assistance.** Please **share with me any concerns or questions** you may have, at any time.

**ANCORA  
IMPARO**

## Accommodations

Accommodations for students with disabilities are determined and approved by Disability Access Services (DAS). If you, as a student, believe you are eligible for accommodations but have not obtained approval please contact DAS immediately at 541-737-4098 or at <http://ds.oregonstate.edu/>. DAS notifies students and faculty members of approved academic accommodations and coordinates implementation of those accommodations. While not required, students and faculty members are encouraged to discuss details of the implementation of individual accommodations.

OSU culture strives toward deepening respect for and understanding of **religious differences** within our community.

## Reach Out for Success

Setbacks can happen. If you encounter difficulties and need assistance, it's important to reach out. Consider discussing the situation with an instructor or academic advisor. Learn about resources for **wellness** and **academic success** at <http://oregonstate.edu/ReachOut>. If you are in **immediate crisis**, please contact the Crisis Text Line by texting OREGON to 741-741 or call the National Suicide Prevention Lifeline at 1-800-273-TALK (8255).

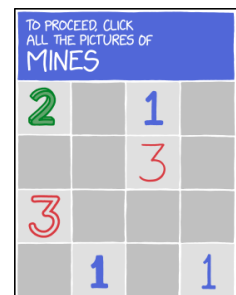
## More Support and Resources

Student Bill of Rights: OSU has twelve established student rights. They include due process in all university disciplinary processes, an equal opportunity to learn, and grading in accordance with the course syllabus: <https://asosu.oregonstate.edu/advocacy/rights> (Psst: want extra credit? Go: <https://forms.gle/7gsbDDaFpW2QftNn7>)

All students are subject to the registration and refund deadlines as stated in the Academic Calendar: <https://registrar.oregonstate.edu/osu-academic-calendar>

## Final Words

I am dedicated to the creation and maintenance of a **positive learning environment**. Be sure you read and understand these guidelines. Most of all - let's have fun as we tackle and solve machine learning challenges together!



xkcd (Randall Munroe)  
<https://xkcd.com/2496/>