AI Club's Project Workshop

- Build your own ML project from scratch
- Add to your portfolio
- Compete for prizes

OSU AI Club

Noah Pragin
Project Manager

A senior majoring in CS with a custom option focusing on robotics and AI. Experience includes TAing for the machine learning course and working on autonomous vehicles.

Lauren Gliane
Project Officer

A 3rd year in Computer Science with a Data Science focus. Joined AI Club to learn AI and help make AI education and projects accessible and welcoming at OSU.

Kellen Sullivan
Project Officer

A senior majoring in CS with a custom AI focus and minoring in math. Experience includes completing multiple AI projects during a data engineering internship.

Ally Chen
Project Officer

A sophomore in CS with a custom focus in AI and Finance, with a Finance minor. Joined AI Club to gain early hands-on experience with AI, an opportunity often limited to upperclassmen.

What We're Building

Goal: Complete ML Project in 10 Weeks

- Work on a guided machine learning project of your choice. No prior experience necessary and we'll help you through every step!

What You'll Gain:

- **Portfolio Project**: A complete ML project to showcase
- **Hands-on Experience**: Real-world data science workflow
- **Resume Material**: Demonstrable AI/ML skills
- **Competition**: Chance to win prizes, AI Club merch and either compute credits or an Amazon gift card!
- **Community**: Work alongside other learners

10-Week Project Timeline Pt. 1

1

Setup and Tutorials

Set up your computer to work on your project and complete tutorials for new tools

2

Problem & Dataset Selection

Define problem, select dataset, identify potential models

3

Data Analysis & Preprocessing

Explore data, create training, validation, and test splits, feature engineering

4

Model Comparison & Tuning

Train 3+ models, basic hyperparameter tuning

10-Week Project Timeline Pt. 2

5

Catch-up / Get ahead

Catch up on anything you missed, polish your model, plan your deployment approach

6

Final Model Selection

Comprehensive tuning and test set evaluation on selected model

7

Model Analysis

Bias detection, interpretability, error analysis

8/9

Deployment Implementation

Build your interface, integrate your model

10

Presentations & Winner Selection

Present your projects and vote on a winner!

Team Structure & Logistics

- ## Team Size
  - Work solo or in groups, whatever works best for you!
  - Solo projects are absolutely doable and encouraged.
  - Groups larger than 3 may have a hard time evenly splitting work
- ## Matchmaking
  - Have friends? Form your own group!
  - Flying solo? We'll help connect you with others!
  - Want to work alone? Perfect!
  - Scan the QR code and fill out the form to help us build teams

Support I

- During Meetings
  - We will walk around during meetings to check in and assist you
  - Recitation-style meetings, where we are happy to give you a hand if you ask

- Outside of Meetings
  - Noah's office hours: 10-11am T+Th in Johnson 121
  - Send questions to Discord!
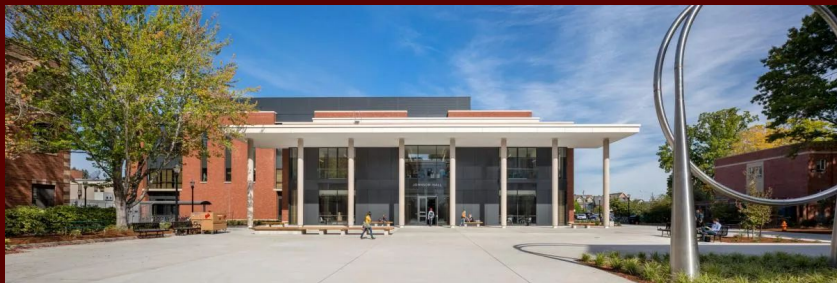
Support II

- The Project Workshop page of the AI Club website has our slides, tutorials, and example project!
  - osu-ai.club/project-workshop

- The first and last slide will always have a QR code to our website
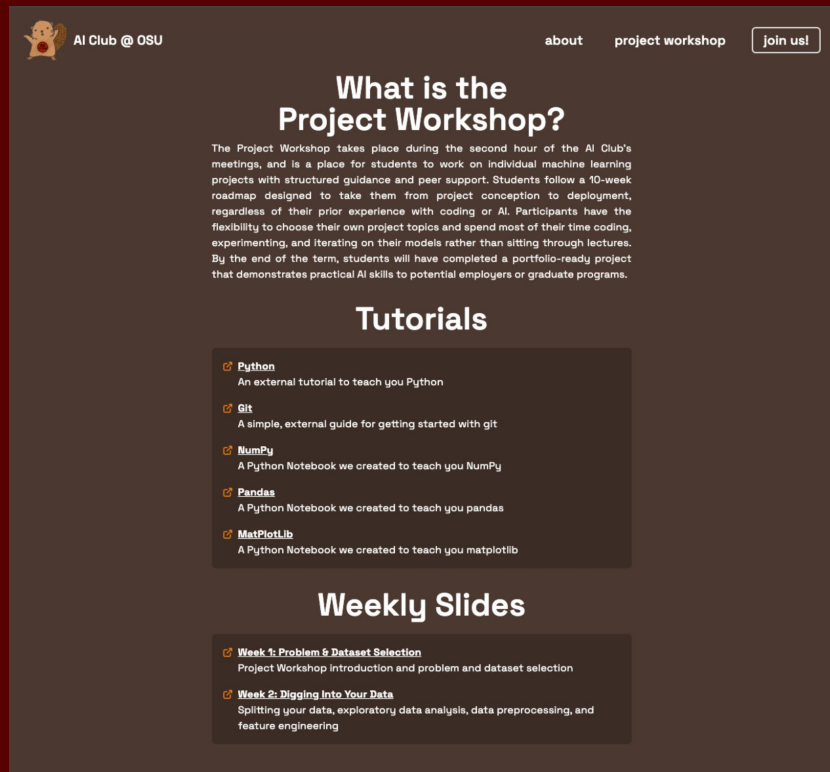
Packages / Tools

- NumPy
  - Numerical processing, used under the hood by everything else
- Pandas
  - Data organization and manipulation, used to aggregate your data before feeding it into a model
- MatPlotLib
  - Plots and visualizations, used to visualize your dataset and model performance
- Scikit-learn
  - Machine learning models and data preprocessing, used for the training and evaluating models

Tips for Success

- Ask questions!
  - ML is a broad topic, and there is no way for us to cover everything

- Find something that interests you!
- Tell us if you see anything wrong or inconsistent!
- Consistency > Intensity

Feedback Survey

- Let us know how today's meeting went!

Week 1: Setting Up & Learning New Skills

- Definition of Done
  - Starting to think about what problem you want to work on
  - Access to ENGR servers
  - Code editor set up
  - Python ready to go
  - Tutorials complete for tools you're unfamiliar with
- Extra Credit
  - Problem identified and defined
  - Dataset selected and accessible

What Makes a Good Problem?

- Clean, available datasets
  - There is no model without data to learn from
- Interesting to you
- Some level of interactivity
- Classification problem?
  - Best suited to beginners, strongest support from us

Limitations of Machine Learning

- Supervised ML Problem Types
  - **Classification**: Is this email spam or not spam?
  - **Regression**: What will the house price be?
- How to Frame a Problem for ML
  - Predict [specific outcome] based on [available data]
  - **Bad**: Detect COVID-19
  - **Good**: Predict COVID-19 severity based on lung CT images
- Guiding Questions
  - Is your prediction target specific?
  - Do you have data to learn from?
 - Can you quantitatively measure success?

Project Ideas & Examples

- Pet Breed Classification
  - **Problem:** Identify cat and dog breeds from photos
  - **Dataset:** Oxford-IIT Pet Dataset
  - **Models:** CNN, ResNet
- Heart Disease Detection
  - **Problem:** Predict heart disease risk from medical data
  - **Dataset:** Cleveland Heart Disease Dataset
  - **Models:** Random Forest, Logistic Regression, SVM
- Network Security
  - **Problem:** Detect malicious network connections
  - **Dataset:** KDD Cup 99, NSL–KDD
  - **Models:** Decision Trees, Neural Networks, Ensemble

Tutorials (Ordered)

- Environment Setup
  - Skim this one even if you're set up, it calls out some good extensions for your code editor
- NumPy
- Pandas
- Scikit-learn
- MatPlotLib?
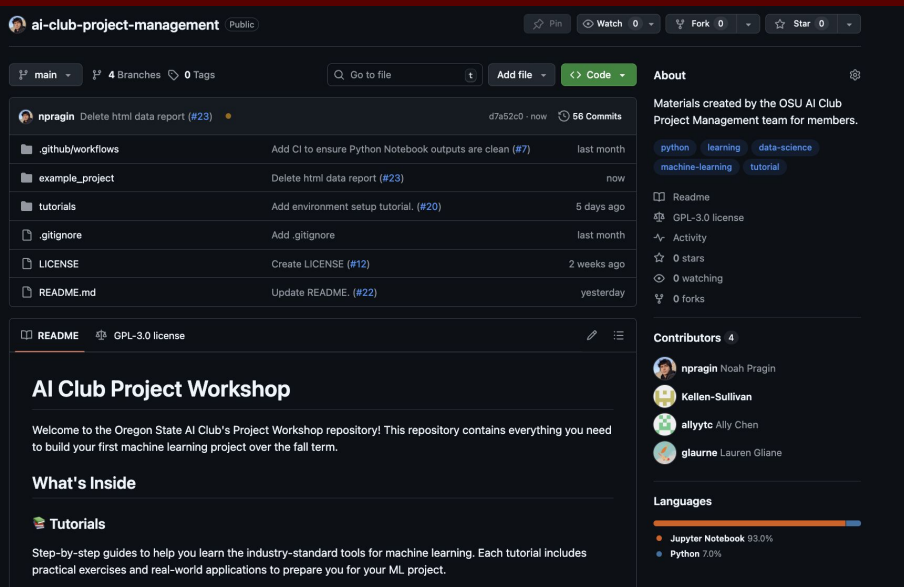
Let's Build Something Amazing!

- Action Items
  - Start thinking about a problem to tackle!
  - Computer set up and ready to go
  - Complete tutorials on unfamiliar tools
- Resources
  - Tutorials on our website; scan the QR code!
  - Project idea slides
- Questions? Stuck?
  - Raise your hand! We're here to help