Model AI Assignments 2022

Todd W. Neller[1], Jazmin Collins[2], Daniel Schneider[3], Yim Register[4], Christopher Brooks[5], Chiawei Tang[6], Chaolin Liu[6], Roozbeh Aliabadi[7], Annabel Hasty[8], Sultan Albarakati[9], Haotian Fang[7], Harvey Yin[7], and Joel Wilson[7]

[1]Gettysburg College, [2]Arcadia University, [3]Code.org, [4]University of Washington, [5]University of San Francisco, [6]National Chengchi University, [7]ReadyAI, [8]Quest Academy, [9]NEOM

Abstract

The Model AI Assignments session seeks to gather and disseminate the best assignment designs of the Artificial Intelligence (AI) Education community. Recognizing that assignments form the core of student learning experience, we here present abstracts of six AI assignments from the 2022 session that are easily adoptable, playfully engaging, and flexible for a variety of instructor needs. Assignment specifications and supporting resources may be found at http://modelai.gettysburg.edu.

FairKalah: Fair Mancala Competition - Todd Neller

Through AI play of the game FairKalah, students learn about game-tree search, alpha-beta pruning, heuristics, feature engineering, machine learning, time management, and related topics. Whereas Mancala presents a barrier for discerning good heuristics because of a significant first-player advantage, FairKalah presents 254 boards where one or two pieces are redistributed from the standard Mancala initial state so as to be fair, i.e. two optimal players draw. Tournament software is provided in Java and Python, as well as a Ludii game definition file to facilitate client flexibility.

A simple minimax player with a simple heuristic is provided. Students are assigned to improve upon this player for real-time competition under time constraints in three primary ways. First, alpha-beta pruning makes the game-tree search more efficient. Second, time management strategies apportion limited computational time wisely in real-time play. Third, the challenge to devise an efficient, helpful heuristic encourages data collection, machine learning, and feature engineering.

While optimal play is computable for FairKalah, adjustable computational time and memory constraints promise to allow greater challenges going forward to push the boundaries of clever, resource-bounded game agents.

Movement and Visual AI - Jazmin Collins

This three-day assignment plan is designed to teach the basic foundations of computer vision and visual AI at the middle school and high school levels. Geared towards those with a foundational understanding of Scratch and block-programming, the assignment utilizes MIT Media Lab's PoseBlocks to create an easily-digestible AI lesson via a Scratch template. Starting off with presentations on movement and the basic concepts of machine learning models, the students will learn about features and training ML models via theoretical diagrams and live demos of browser-based AI. The presentations focus on interaction with the students and encourage student feedback, questions, and interaction with the lessons being led by the instructor. Since this assignment is also geared to be used in either an in-person or a virtual format, the dependencies of the project are quite low, using only browser-based methods such as Google's Teachable Machine to quickly create and launch custom ML models based on webcam data. These models give the students hands-on experience with the creation of a basic AI. Finally, they can integrate these models and other pre-created models into Scratch applications to understand the incorporation of ML models into meaningful AI applications based on computer vision.

When Your Neighbor Is a Zombie: Zombie KNN - Daniel Schneider and Yim Register

A powerful, interpretable, and multipurpose AI algorithm is k-nearest neighbors (KNN). KNN capitalizes off of similarities between data points for either regression or classification. KNN is overwhelmingly taught using abstract data like red triangles and blue squares on a 2D plane. We present a completely unplugged KNN lesson that engages its K12 audience in a thrilling adventure to cross a map to safety during the zombie apocalypse. Using a dataset that characterizes how many zombies congregate in certain areas (using features like noise level), students use the information to infer which areas on the map are likely to have fewer zombies and grant them safe passage. Students learn to make a prediction of how many zombies are likely in an area by tallying up the number of similarities between the new location and locations in their dataset. Next, they learn to take the mean number of zombies from those similar locations, and use this as a prediction. Finally, the lesson engages students with ideas about model performance in high-stakes scenarios. While the zombie game is just for fun, we prompt students to also think about model accuracy for problems of

distributing resources, gauging real world infection rates, or health decisions.

Reflecting on Bias - Christopher Brooks

This set of assignments provides students with the opportunity to read about and reflect on the many different ways in which bias can occur in AI and ML problems. Students will learn about the technical definition of bias, how this relates to the more colloquial version, why it is so difficult to define what it means to be fair, and the importance of understanding what biases exist in real-world data sets. Also included are strategies and resources for class discussion, connecting to technical topics, and a discussion of specifications grading as an approach for assigning written work.

Projecting Your Data - Chiawei Tang and Chaolin Liu

In this assignment, we will use Pytorch to implement a Convolutional Neural Network (CNN) classifier for the MNIST datasets and discuss how to use Pytorch to build a custom dataset and design the model. After finishing the classifier, we will build an embedding projector based on the model to observe the behavior of the model on the dataset. This assignment will combine the concepts including classification, convolution neural network, and working with embeddings.

Different hyperparameters or model architectures can lead to different outputs on the same dataset, and different dimensionality reduction algorithms can change the way we project the embeddings into a 2D or 3D space, which in turn affects the way we interpret the model. After building the embedding projector, we will explore how the projection of the embeddings changes under different settings and analyze the behavior of the model.

This assignment is organized as follows:

- Part A. Data exploration
- Part B. Building the classifier
- Part C. Building the embedding projector
- Part D. Analyzing the embeddings

Much of this assignment is inspired by Google's embedding projector (Smilkov et al. 2016), and we will explore how to make and improve our own embedding projector based on our needs.

Introduction to Problem-Solving With Data: A Fresh Squeeze on Data - Roozbeh Aliabadi, Annabel Hasty, Sultan Albarakati, Haotian Fang, Harvey Yin and Joel Wilson

Data bias arises when the available data does not represent the entire study population, often due to a small sample set. To facilitate a comprehensive introduction to data bias, this assignment provides an authentic way for students from third to fifth grade to explore and experiment with data bias. The 90-minute session in this assignment includes two main activities and requires no prerequisite knowledge or dependencies. The teacher and students work together to collect data and explore the relationship between data and bias in multimedia teacher resources, including a visual book, videos, web tools, and interactive quizzes. This assignment encourages students to examine data bias's practical, real-life implications in a captivating story with vibrant characters and real-world problems. To achieve such AI educational goals, the storyline revolves around a young female character who successfully runs a lemonade stand with the help of her data scientist mother and a diverse group of friends. The story and lesson plans could be redesigned for high school or college level students utilizing the worksheet and online resources. Understanding data bias requires intensive study of statistics, machine learning, and data analytics with subsequent coursework.

References

Smilkov, D.; Thorat, N.; Nicholson, C.; Reif, E.; Viégas, F. B.; and Wattenberg, M. 2016. Embedding Projector: Interactive Visualization and Interpretation of Embeddings. arXiv:1611.05469.