(Will be updated for Spring 2024)

95-891: Introduction to Artificial Intelligence

Fall 2023 schedule (12 units):

- Lectures: TTh 5:00 PM - 6:20 PM Eastern, Hamburg Hall A301 (or on Zoom as indicated below)
- Recitations on every other Friday from 9:30 AM - 10:50 AM in Hamburg Hall 1002

Instructor: David Steier (steier@andrew.cmu.edu)
Teaching Assistant:

- Xiaobin Shen (xiaobins@andrew.cmu.edu)

Driven by the combination of increased access to data, computational power, and improved sensors and algorithms, artificial intelligence (AI) technologies are entering the mainstream of technological innovation. These technologies include search, machine learning, natural language processing, robotics and image processing.

The course begins by describing what the latest generation of artificial intelligence techniques can actually do. After an introduction of some basic concepts and techniques, the course illustrates both the potential and current limitations of these techniques with examples from a variety of applications. We spend some time on understanding the strengths and weaknesses of human decision-making and learning, specifically in combination with AI systems and on ethical and policy implications of new AI capabilities. Exercises will include hands-on application of basic AI techniques as well as selection of appropriate technologies for a given problem and anticipation of design implications. In a final project, groups of students will participate in the creation of an AI-based application.

Course Learning Outcomes

The main learning objectives of the course are to:

1. Identify problems where artificial intelligence techniques are applicable
2. Apply selected basic AI techniques; judge applicability of more advanced techniques.
3. Participate in the design of systems that act intelligently and learn from experience.

Course Prerequisites

This course is primarily aimed at students with technical backgrounds who wish to design and develop products and services using AI. A background in basic statistics is required for the course. Students need at least a basic knowledge of Python to complete the assignments for this course. Students who have not taken 90-812 or 95-888 or have equivalent background will be required to complete supplementary work to learn Python at the beginning of the course.

Instructor: David Steier (PhD, CMU SCS '89)

David Steier joined the CMU faculty in 2018 as Distinguished Service Professor in the Heinz College School of Information Systems and Management. Prior to joining CMU, David was Managing Director in Deloitte Consulting's Data Science practice. At Deloitte, David helped clients use advanced data analytics and visualization in a variety of industries. Prior to Deloitte, David was Director in the Center for Advanced Research at PwC, Senior Director of Technology and Business Development at Kanisa, and Managing Director at Scient. David holds a Ph.D. in computer science from Carnegie Mellon and a bachelor's degree in computer science from Purdue University.

Course Resources and Policies

Canvas and Piazza

An online site with this syllabus, readings, and other resources has been created in Canvas at https://canvas.cmu.edu/courses/36428 . The system is highly catered to getting you help fast and efficiently from classmates, the TAs, and myself. Rather than emailing questions to the teaching staff, we encourage you to post your questions on Piazza. Find our class signup link at https://piazza.com/cmu/fall2023/95891a .

Text

The primary reference for the course is Russell, S. & Norvig, P. *Artificial Intelligence: A Modern Approach*, Pearson, 2020. This is the fourth edition of the leading textbook in AI, generally accepted as the most comprehensive reference on the subject. It is a substantial update to the third edition, so investing in the latest edition is worthwhile, especially if you plan to do further work in AI. One copy is also available on virtual reserve through the CMU library. This text will be supplemented by on-line material as listed in the course outline below.

Homework, final projects and grading

There will be 6 homework assignments each due at midnight (Eastern Time) per the following schedule:

	Assignment	**Due**
1)	Search	Sep 14
2)	Classification and clustering	Oct 5

3) Computer vision Oct 26
4) Natural language Nov 9
5) Fairness in AI Nov 21
6) Artificial General Intelligence Nov 30

Each assignment will count for 10 percent of the grade, with the lowest grade dropped, for a total of 50 percent. Late assignments (without a written excuse for medical/family/etc. emergencies) will be penalized at the rate of 10% of the assignment's grade per day late. A final project presentation and report will count for 20 percent of the grade. The final projects are meant to be done in groups, which should split the work equally and report who did what in the final report. Three closed-book quizzes on Sep 14, Oct 12, and Nov 14 will count for another 10 percent of the grade. There is no final exam for this class. The remaining 20 percent will be based on class attendance and participation in discussions. Two absences are permitted, with further absences causing a proportional deduction in the class participation grade. When the course session is in-person, in-person attendance is required for attendance credit, unless there are medical circumstances requiring remote attendance. When you must miss class, please notify me (at least 24 hours in advance except for illness/emergency), so that we can discuss alternative arrangements for catching up on class and associated work. If you encounter extenuating circumstances and must miss more than two classes, please come and discuss the issue with me; I would like to find a way to support you.

Grading will be on a straight scale as follows (with no rounding up):

A+	98.0-100%	B+	88.0-89.9%	C+	78.0-79.9%
A	92.0-97.9%	B	82.0-87.9%	C	72.0-77.9%
A-	90.0-91.9%	B-	80.0-81.9%	C-	70.0-71.9%

Everyone taking the class should register for a letter grade. Auditing the class, or taking the class Pass/Fail, is intended for extremely rare circumstances and only with consent of the instructor.

Course Outline

This is a full-semester course, planned in general around two sessions per week of 1 hour 20 minutes apiece. Sixteen of the lecture sessions will be held in Hamburg Hall A301, and the remaining nine sessions will be held on Zoom as indicated below. Recitations will be held in Hamburg 1002. Note that there are no classes on October 17 and 19 due to Fall Break, nor on November 23 due to the Thanksgiving holiday.

- Week 1: Introduction to AI (Lectures Aug 29 & 31, Hamburg A301)
  - **Topics**
    - Introductions
    - Course structure and policies
    - What is AI (using case studies)
    - History of AI
    - Proposing and evaluating AI applications
  - **Readings**

    - Russell & Norvig, Chapter 1, "Introduction" in *Artificial Intelligence: A Modern Approach*, 2020
    - (optional) Leviathan Y, and Matias, Y, "Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone," Google AI Blog, May 8, 2018 https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html
- **Assignment out**: HW #1

- Week 2: Search (Lectures Sep 5 & 7, Hamburg A301; Recitation Hamburg 1002)
    - **Topics**
        - Importance of search for AI
        - Uninformed and informed search
        - Adversarial search
        - Local search (gradient descent)
        - Recitation topic: Implementing search
    - **Readings**
        - "Chapters 3: Solving Problems by Searching," and "Chapter 5: Adversarial Search" in Russell & Norvig, *Artificial Intelligence: A Modern Approach*, 2020

- Week 3: Reasoning with Uncertainty (Lectures Sep 12 & 14, Hamburg A301)
    - **Topics**
        - Uncertainty
        - Bayesian networks
    - **Readings** in Russell & Norvig, *Artificial Intelligence: A Modern Approach*, 2020
        - Quantifying uncertainty: Ch. 12
        - Probabilistic reasoning: Ch. 13.1-13.3
    - **Assignment due**: HW# 1 (Sep 14)
    - **Quiz #1** (Sep 14)

- Week 4: Introduction to Machine Learning (Lectures Sep 19 & 21, Hamburg A301; Recitation Sep 22, Hamburg 1002)
    - **Topics**
        - What is machine learning?
        - Supervised vs. unsupervised learning
        - Regression -- linear, logistic, ridge
        - Classification – decision trees, SVM, random forests
        - Model evaluation
        - Dimensionality reduction: PCA
        - Clustering – k-means, hierarchical clustering

        - Recitation topic: Implementing machine learning
    - **Readings**
        - "Chapter 19: Learning from Examples" in Russell & Norvig, *Artificial Intelligence: A Modern Approach*, 2020
        - C. Aggarwal, Chapter 9 "Unsupervised Learning" in *Artificial Intelligence: A Textbook*, 2021, Springer, file available on Canvas
        - Huneycutt, J., "An introduction to clustering algorithms in Python", May 29, 2018, https://towardsdatascience.com/an-introduction-to-clustering-algorithms-in-python-123438574097
    - **Assignment out**: HW #2

- Week 5: Machine learning (continued) (Sep 26 & 28, Hamburg A301)
    - **Topics**
        - Semi-supervised learning
        - Expectation maximization
        - Reinforcement learning
    - **Readings**
        - Russell & Norvig, Chapters 17.1 "Sequential Decision Problems", 20 and 22, "Learning Probabilistic Models" and "Reinforcement Learning" in *Artificial Intelligence: A Modern Approach*, 2020
        - (optional) van Engelen, J.E., Hoos, H.H. A survey on semi-supervised learning. Mach Learn 109, 373–440 (2020). https://link.springer.com/article/10.1007/s10994-019-05855-6

- Week 6: Deep Learning (Oct 3 & 5, Hamburg A301; Recitation Oct 6, Hamburg 1002)
    - **Topics**
        - Neural networks and back-propagation
        - Convolutional neural networks
        - Recurrent neural networks and LSTMs
        - Transfer learning
        - Recitation: Implementing deep learning in PyTorch
    - **Readings:**
        - Russell & Norvig, Chapter 21, "Deep Learning" in *Artificial Intelligence: A Modern Approach*, 2020
        - (optional) 3Blue1Brown, "But what is a neural network," Chapter 1 Deep learning," 2017 (20 min video) https://www.youtube.com/watch?v=aircAruvnKk&vl=en
    - **Assignment due**: HW #2 (Oct 5)
    - **Assignment out**: HW #3

- Week 7: Computer Vision (Oct 10 & 12, Hamburg A301)
    - **Topics**

    - Introduction to computer vision
    - Image segmentation
    - Edge and motion detection
    - Object classification
    - Pre-trained vision models
- **Readings**
    - Russell and Norvig, Chapter 25, "Computer Vision" in *Artificial Intelligence: A Modern Approach*, 2020
    - PyTorch Dataloader video: https://www.youtube.com/watch?v=zN49HdDxHi8
    - TensorFlow, "Image Recognition", July 30, 2018, https://www.tensorflow.org/tutorials/images/image_recognition
    - He, K,Dec., 2015, "Deep Residual Learning for Image Recognition", https://arxiv.org/pdf/1512.03385.pdf
- **Quiz #2** (Oct 12)
- **Assignment due**: Final project groups formed and initial proposal (Oct 10)

- Week 8: Natural Language Understanding (Lectures Oct 24 & 26, Zoom; Recitation Oct 27, Hamburg 1002)
    - **Topics**
        - Intro to natural language understanding
        - Machine translation
        - Sentiment analysis
        - Language models for natural language understanding
        - Recitation: Implementing transformers in PyTorch
    - **Readings**
        - Russell & Norvig, "Chapter 23: Natural Language Processing" and "Chapter 24: Deep Learning for Natural Language Processing," in *Artificial Intelligence: A Modern Approach*, 2020
        - (Optional) Brown, T.B. et al, "Language Models are Few Shot Learners" (GPT-3), May 28, 2020, https://arxiv.org/pdf/2005.14165.pdf
        - (optional) Devlin, J. et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", 24 May 2019, https://arxiv.org/pdf/1810.04805.pdf
    - **Assignment due**: HW #3 (Oct 26)
    - **Assignment out**: HW #4

- Week 9: Speech and Natural Language Interaction (Lectures Oct 31 & Nov 2, Zoom)
    - **Topics**
        - Speech recognition
        - Speech synthesis
        - Natural language generation

    - Language models for natural language interaction
    - Case studies: Google Duplex and ChatGPT
- **Readings**
    - OpenAI, "ChatGPT: Optimizing Language Models for Dialog" November 30, 2022, https://openai.com/blog/chatgpt/
    - (optional) Ouyang, L. "Training language models to follow instructions with human feedback", 4 Mar 2022, https://arxiv.org/pdf/2203.02155.pdf

- Week 10: Ethics & AI; Introduction to Robotics (Lectures Nov 7 & Nov 9, Zoom; Recitation Hamburg 1002)
    - **Topics**
        - Privacy
        - AI and the future of work
        - Algorithmic bias
        - Introduction to robotics
        - Navigation and path planning
        - Learning and robotics: Reinforcement learning
        - Recitation: Tools for exploring fairness
    - **Readings**
        - Russell & Norvig, "Chapter 27: Philosophy, Safety and Ethics of AI" in *Artificial Intelligence: A Modern Approach*, 2020
        - E. Ntoutsi, et. al., "Bias in data-driven artificial intelligence systems—An introductory survey," Wiley Online Library, 03 February 2020, https://onlinelibrary.wiley.com/doi/full/10.1002/widm.1356
        - Russell & Norvig, "Chapter 26.1-26.7: Robotics" in *Artificial Intelligence: A Modern Approach*, 2020
        - (optional) Adam Savage, "How Boston Dynamics' Spot Robot Works," Apr 8, 2020, (19 minute video), https://www.youtube.com/watch?v=R-PdPtqw78k&t=896s
        - (optional) Mnih, et al, Dec 2013, "Playing Atari with Deep Reinforcement Learning," https://arxiv.org/pdf/1312.5602.pdf
    - **Assignment due**: HW #4 (Nov 9)
    - **Assignment out**: HW #5

- Week 11: Robotic control and autonomous vehicles (Lectures Nov 14 & 16, Hamburg A301)
    - **Topics**
        - Robotic control
        - Human-robot interaction
        - Autonomous vehicles
    - **Readings**

        - Russell & Norvig, "Chapter 26.8 Humans and Robots" in *Artificial Intelligence: A Modern Approach*, 2020
        - L. Fridman, "Human-Centered Autonomous Vehicle Systems: Principles of Effective Shared Autonomy", 3 Oct 2018, https://arxiv.org/pdf/1810.01835.pdf
    - **Quiz #3** (Nov 14)

- Week 12: Artificial General Intelligence (Lecture Nov 21 on Zoom)
    - **Topics**
        - Artificial General Intelligence
    - **Readings**
        - J. Reed, et. Al, "A Generalist Agent", May 19, 2022, https://arxiv.org/pdf/2205.06175.pdf
    - **Assignment due**: HW #5 (Nov 21)
    - **Assignment out**: HW #6

- Week 13: Infrastructure for AI; Futures of AI (Lecture Nov 28 & 30, on Zoom)
    - **Topics**
        - Parallel and distributed computing for scalability
        - MLOps
        - Emerging technologies for large language models
        - Emerging developments such as brain-computer interfaces
    - **Readings**
        - https://a16z.com/2020/10/15/the-emerging-architectures-for-modern-data-infrastructure/
        - Google Cloud, MLOps: Continuous delivery and automation pipelines in machine learning, https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning
        - Russell & Norvig, "Chapter 28: The Future of AI", in *Artificial Intelligence: A Modern Approach*, 2020
    - **Assignment due**:
        - HW #6 (Nov 30)

- Week 14: Final project presentations (Dec 5 & 7, held on Zoom)
    - **Topics**
        - Final presentations
    - **Assignment due**:
        - Final project presentations
        - Final report (Dec 13)

Academic Integrity

Students are expected to strictly follow Carnegie Mellon University rules of academic integrity in this course. This means in particular that unless otherwise specified, homework are to be the work of the individual student using only permitted material and without any cooperation of other students or third parties. It also means that usage of work by others is only permitted in the form of quotations and any such quotation must be distinctively marked to enable identification of the student's own work and own ideas. All external sources used must be properly cited, including author name(s), publication title, year of publication, and a complete reference needed for retrieval. The same work may not be submitted for credit in multiple courses. Violations will be penalized to the full extent mandated by the CMU policies. There will be no exceptions.

Certain assignments in this course will permit or even encourage the use of generative artificial intelligence (AI) tools, such as ChatGPT. When AI use is permissible, it will be clearly stated in the assignment prompt posted in Canvas. Otherwise, the default is that use of generative AI is disallowed. In assignments where generative AI tools are allowed, their use must be appropriately acknowledged and cited. For instance, if you generated the whole document through ChatGPT and edited it for accuracy, your submitted work would need to include a note such as "I generated this work through Chat GPT and edited the content for accuracy." Paraphrasing or quoting smaller samples of AI generated content must be appropriately acknowledged and cited, following the guidelines established by the APA Style Guide. It is each student's responsibility to assess the validity and applicability of any AI output that is submitted. You may not earn full credit if inaccurate on invalid information is found in your work. Deviations from the guidelines above will be considered violations of CMU's academic integrity policy. Note that expectations for "plagiarism, cheating, and acceptable assistance" on student work may vary across your courses and instructors. Please email me if you have questions regarding what is permissible and not for a particular course or assignment.

Diversity

It is my intent that students from all diverse backgrounds and perspectives be well served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity: gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture. Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally or for other students or student groups.

Disability Accommodations

If you have a disability and have an accommodations letter from the Disability Resources office, I encourage you to discuss your accommodations and needs with me as early in the semester as possible. I will work with you to ensure that accommodations are provided as appropriate. If you

suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, I encourage you to contact them at access@andrew.cmu.edu.

Mental Health

As a student, you may experience a range of challenges that can interfere with learning, such as strained relationships, increased anxiety, substance use, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may diminish your academic performance and/or reduce your ability to participate in daily activities. CMU services are available, and treatment does work. You can learn more about confidential mental health services available on campus at: http://www.cmu.edu/counseling/. Support is always available (24/7) from Counseling and Psychological Services: 412-268-2922.

