CAP 4630 - Introduction to Artificial Intelligence (Spring 2025)
Project 4 - Due Date: Friday, 4/25/2025 11:59 pm

In this final project, you will form a team, of **up to 3** people, to address a **machine learning** problem of your own choice. Some problems are suggested to you to select from, but you are not limited to this pool provided.

1 Important Dates

- Proposal submission due: 11:59pm, Friday, Apr 11
- Progress report email due: 11:59pm, Friday, Apr 18
- Final Project submission due: 11:59pm, Friday, Apr 25

See the following Guidelines sections for detailed tasks and requirements.

2 Project Guidelines

To complete this project you may form a team of up to 3 people. You must choose a topic for this project that solves a **machine learning** problem (e.g., classification, regression, or clustering). Your team may take one of the two approaches to formulate your project:

1. **Research**: Pick a dataset, identify a machine learning problem investigating the dataset, implement algorithms to learn models to address the problem, and present and analyze the results.
2. **Application**: Design and implement an innovative application using existing machine learning algorithms and techniques.

If you choose to work on a research project, make sure novelty in either the problem itself or the algorithms to solve it, and you may choose two different learning models (could be of the same type) and perform comparative analysis of them.

If you choose to work on an application project, make sure novelty in the objective, design, or implementation of your system.

2.1 Suggested Problems

1. Image classification for CIFAR-10/CIFAR-100: predict classes of images in a dataset of 60 thousand 32x32 colored images in 10/100 classes. This dataset is accessible at `https://www.cs.toronto.edu/~kriz/cifar.html`.
2. GTMNERR marsh grass species classification: predict species of marsh grass in 33x33 colored images taken in our local Guana Tolomato Matanzas marine sanctuary. This dataset is accessible at `http://unfail.ccec.unf.edu/marshdata.html`.

3. Recyclable material classification: predict types of recycle materials in 150x150 colored images. This dataset is accessible at `http://unfail.ccec.unf.edu/recycledata.html`.
4. Movie review predictions: movie review analytics using Rotten Tomatoes movies and critic reviews dataset. This dataset is accessible at `https://www.kaggle.com/datasets/stefanoleone992/rotten-tomatoes-movies-and-critic-reviews-da`
5. Smart personal agent that integrates GPT power to assist personal tasks: traveling, scheduling, education, etc. Check out tutorials on using OpenAI's GPT API at `https://platform.openai.com/docs/tutorials`.

3 Proposal Submission Guidelines

A PDF proposal must be submitted on Canvas by Friday, Apr 11. The proposal should contain the following information about your project:

1. Names of team members (up to 3 people).
2. The topic with clear statement on whether research or application.
3. Problem statement if research project, or system statement if application project.
4. The planned activities and time line detailing who will do what by when: this must include every student in the team and every one should take on relatively same work load.

4 Progress Report Email Guidelines

Only one email must be sent to me by Friday, Apr 18, from the team coordinator, with other team members copied. This email should contain the following information:

1. What your team has done, both team wise and member wise
2. Show case what has been done (e.g., screenshots)
3. What your team plans to do, both team wise and member wise
4. Questions if any

5 Final Report Guidelines

For your convenience, the final project report will be written using the AAAI publishing template: `https://aaai.org/conference/aaai/aaai-25/submission-instructions/`. You will click on the "AAAI-25 Author Kit" button once on that website to download the author kit. In the author kit, you will go to CameraReady folder and find the AAAI template in both Word and LaTex. Feel free to choose the one you feel comfortable with. The submitted report should be a PDF produced from the Word or LaTex template.

Please make sure that your report be **3-4 pages, excluding references section**.

Regarding the writing,

1. be sure your report flows well, is clear, and makes sense.
2. be concise and make sure that each paragraph/section presents meaningful information.
3. points will be taken for grammatical and spelling errors.

Similar to research papers, the report should be organized by the following sections:

1. Abstract: a brief description of what is in the report
2. Introduction or motivation: why do you conduct this project and why is it important?
3. Related work: existing research relevant to your project

4. Methodologies/algorithms/approaches
5. Experiments and results
6. Conclusions
7. References

However, the sections can vary depending on the nature of the project you conduct.

6 Final Project Submission Guidelines

Submit a zip file of name [your-team-last-names]_Project4.zip of the following items to Canvas by Apr 25.

1. Proposal (due Apr 11)
2. Final report
3. Individual contributions report
   (a) This should provide an overall percentage distribution of team members over the entire project work, and a brief justification thereof.
   (b) If this report is missing, all will be treated the same fair share percentage ($\frac{1}{N}$ assuming N people in the team).
4. Source code: zip the following to Programs.zip
   (a) A directory that contains all your Python programs that are may be .py or .ipynb.
   (b) A README file that contains instructions to run your Python programs.

7 Grading

Total: 100 pts

1. Proposal (20 pts)
2. Progress Email (10 pts)
3. Final Report (20 pts)
4. Source Code (40 pts)
5. Individual Contribution (10 pts)

Note that the first 4 criteria are team rubrics that apply to all team members and the last criteron "Individual Contribution" is per every student. The details for this criterion is as follows.

Let N be the size of a team, and X be the individual contribution in terms of percentile of a student in such team, then the student will get 0 points if $X < 50\% \cdot \frac{1}{N}$, 2 points if $50\% \cdot \frac{1}{N} \leq X < 60\% \cdot \frac{1}{N}$, 4 points if $60\% \cdot \frac{1}{N} \leq X < 70\% \cdot \frac{1}{N}$, 6 points if $70\% \cdot \frac{1}{N} \leq X < 80\% \cdot \frac{1}{N}$, 8 points if $80\% \cdot \frac{1}{N} \leq X < 90\% \cdot \frac{1}{N}$, and 10 points if $X \geq 90\% \cdot \frac{1}{N}$.

Let's consider a team of 3 students A, B and C. Say in the individual contributions report, the individual contributions are 40%, 35%, and 25% for A, B and C respectively. Then, per the "Individual Contribution" rubric, A and B will get 10 points and C will get 6 points.

8 General Guidelines

Academic integrity is defined in the syllabus.

EXCEPTION: With proper references, you may refer to open source programs that accumulate up to 20% of your programs in the project.