Final Project

Please go here to enter your project name and team members (also posted on our website)

<https://docs.google.com/spreadsheets/d/1-bGTvAsl-N_n1SVzoZa8O0FJTM3HkJWR7hPM4jOQ6XA/edit?usp=sharing>

**Deadlines:**

* Final presentation: 1:00 pm, December 6, 2021
* Zoom Link: <https://us02web.zoom.us/j/88041556245?pwd=ZUhWaVNqcjQzclYwZUFqUkdHaGtzQT09>
* Paper/code/slide submission: 11:59 pm, December 6, 2021

**Overview:**

* The goal of the final project is to teach you how to identify interesting ideas, conduct research, and write a scientific paper.
* Each group will select an interesting research question, a machine learning algorithm (within or outside of this class’s curriculum), and a public real-world dataset.
* The students will come up with a list of experiments they want to conduct to answer the question. The instructor (Dr. Hien Van Nguyen) and the teaching assistant (Liqiang Huang) will be available to assist students throughout this process if they have questions.
* If the final paper’s quality is good enough, you may submit it to a reputable journal (impact factor > 4.0, top 10% of journals), and the instructor will pay the publication fee.

**Group size:** From 5 to 7 students

**Examples of algorithms outside of the class’s curriculum:**

* Supervised learning
  + Advanced deep neural networks for image detection, segmentation, and classification
  + Capsule networks
  + Bayesian deep networks
  + Recurrent neural networks for modeling time-series data
  + Open-set deep networks
  + Attention-augmented deep networks
* Unsupervised learning
  + Advanced clustering and segmentation algorithms
  + Generative adversarial networks
  + Variational autoencoder networks
* Visualization and proofreading
  + Methods for visualizing important features captured by machine learning algorithms
  + Methods for visualizing error patterns made by your classifiers
  + Methods for efficient correction/editing of prediction mistakes

**Comprehensive list of public datasets:**

* <https://github.com/beamandrew/medical-data>
* <https://github.com/adalca/medical-datasets>
* <https://grand-challenge.org/challenges/>
* <https://www.kaggle.com/datasets>
* <https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research>

**Evaluation Criteria:**

At the end of the project, each group will present the research question, literature survey, methodology, key results, and challenges they have to overcome during the execution. The presentation will be graded by a panel of 3 machine learning researchers, including the instructor and two machine learning Ph.D. students. The evaluation will be based on whether:

* the research question is interesting (20%)
* the literature review is thorough (20%)
* the experimental methodology is rigorous (30%)
* the discussion of results is insightful (30%)

To ensure that all members contribute to the project, group members will submit confidential reviews of other group members. For example, John’s groupmates will answer the question “Did John contribute substantially to your project?” If John receives less than 50% of positive responses, his score will be 30% lower than the corresponding group’s score.

Students will submit their presentation slides, paper, and code to the blackboard website at the end of the semester. Your paper should have a maximum of 8 pages (excluding references) and follow IEEE template.

**IEEE Template:**

You can find MS Word and Latex templates of IEEE papers in the following links: [https://journals.ieeeauthorcenter.ieee.org/create-your-ieee-journal-article/authoringtools-and-templates/ieee-article-templates/templates-for-transactions/](https://journals.ieeeauthorcenter.ieee.org/create-your-ieee-journal-article/authoring-tools-and-templates/ieee-article-templates/templates-for-transactions/)