PROJECT DESCRIPTION

The project is an essential part of this class. It will allow you to demonstrate your Machine Learning (ML) skills and create something that you are proud of. It can also be a valuable addition to your projects portfolio that you can demonstrate to prospective employers.

Project Requirements

- The project has two key components to it:
  - Understanding a recent machine learning technique and associated algorithm(s)
  - Apply it to a standard dataset of sufficient complexity. **You have to code the main part of the algorithm without using any built-in library.** You can use libraries for pre-processing, loading, analysis of results, etc.

- A list of acceptable techniques/algorithms is mentioned later in this document. You are also free to propose any additional technique to the instructor. Note that you have to wait for approval before starting on a proposed technique.

- Your project deliverables should consist of two major parts:

  1. A report in IEEE conference format
     https://www.ieee.org/conferences/publishing/templates.html

     It should have the following sections at the minimum:
     - Abstract
     - Introduction and background work
     - Theoretical and conceptual study of the technique/algorithm you would like to implement
     - Results and analysis. Please include results in tabular or graphical formats. Be sure to analyze your results well.
     - Conclusion and future work
     - References

The report excluding the references should be 4-6 pages long. The final file should be converted to PDF format before submission.

2. Your code, link to dataset, results, and instructions for compiling and running

Below are the requirements:

- You have to code one or more algorithms from the technique that you have studied on your chosen dataset without using any pre-built library. You are free to use pre-processing and evaluation libraries, but the main algorithm should be coded by you.

- The coding can be done in any language of your choice. Be sure to include instructions on how to compile and run your code.

- You are free to choose a dataset of your choice from sources like Kaggle, or any other source. Do not include the dataset as part of the deliverables, instead host it on your UTD web account. This will allow the TA to run your code without having to search the dataset or download huge files. If you do not know how to host data on UTD account, contact the TA.

- A log file of your experiments (i.e. runs) and parameters should be maintained and submitted. Example of a log file can be:

Experiment Number	Parameters Chosen	Results
1	Neural Net: Number of layers = 4 Neurons = (8, 8, 4, 2) Error Function = RMSE Regularization Parameter = 0.6	Train/Test Split = 80:20 Size of dataset = 10,000 Training Accuracy = 95% Test Accuracy = 88% Training RMSE = 1.67 Test RMSE = 3.08 .....
2	.....	.....
....	.....	.....

- Below are some further administrative requirements:

    - All contents of your report must be original. You have to write the report in your own words. It is acceptable to include figures from the references, provided you state the source clearly in the caption.

    - Your report will be checked for plagiarism. Any violation will carry strong penalties, including reporting the incident to university authorities.

    - Team size requirements: Project can be done in teams of 1 to 4 students. More than 4 students cannot be in a team under any circumstances. You can only form team within the same class and section. You are not allowed to work or collaborate with students from other sections.

Project Topics

Below is the list of topics that you can choose from:

- Deep Reinforcement Learning
- Recurrent Neural Networks (RNN) for machine translation
- Recurrent Neural Networks (RNN) for time series prediction (e.g. stock market, weather, hurricane intensity data)
- Image and video captioning with deep neural networks
- Autoencoders for bioinformatics or image processing
- Scene recognition with deep neural networks
- Genetic sequence analysis using deep neural networks
- Reinforcement learning for game playing
- Meta-Learning
- Transfer Learning
- Adversarial Machine Learning
- Statistical Relational Learning
- Human assisted Machine Learning

Deliverables and Deadlines

Deadline	Project Phase	Deliverable
Monday October 25 Midnight	Project Status Report	Submit a one-page report containing following on eLearning: • Project Topic • Team Members • Technique/Algorithm you plan to implement • Dataset details, such as number of features, instances, data distribution • Coding language / technique to be used • Preliminary Results (if available)
Monday November 29 Midnight	Final Report	• Complete project deliverables as described in the requirements above to be submitted via eLearning ** Your report and code will be checked for plagiarism **