MSSE 277B: Machine Learning Algorithms

Project Assignment

Assigned March 20 and Due April 28

For the final project we will develop a supervised learning ANN model applied to the ANI-1 data set, and also have you consider at least one additional unsupervised/supervised method as well. The finals project has the following expectations for assessment:

a. demonstrate best practices in regards data preparation and analysis of data before starting any subsequent supervised learning strategy; this should be evident in the reporting of results.
b. An individual jupyter notebook should be maintained during the course of the project to show your process of thinking and experimenting.
c. Report the result on your final production run and comment on whether your goal is achieved in the end.

We will do a check-in once per week to see steady progress with appropriate entries of dates in the jupyter notebooks on what has been accomplished. I.e. this is not a project assignment that should be finished the night before. This will be part of the assessment.

Grading breakdown:

(1) Consulting with Prof. Head-Gordon and Zi-Yuan about ideas for an additional ML method which should be finalized by March 24 (10% of the grade)
(2) Progress report due April 6. (10% of the grade) You'll need to submit a one page writeup of the problem and progress to solve it. Include the goal of your project, description on the dataset and the plans for your deep learning explorations.
(3) Progress check-in due April 13. (10% of the grade) Submit your jupyter notebook to gradescope, and also Ziyuan will check your code progress in your final project repository. At this point you should have a working code well on its way that can train the network, demonstrated on at least subset of the data.
(4) Progress check-in due April 20. (10% of the grade) Submit your jupyter notebook to gradescope, and also Ziyuan will check your code progress in your final project repository. At this point your code and solutions should be in production phase.
(5) Final presentation on April 24. (40% of the grade) This is expected to be a 10 minute presentation to the class;
(6) Final form of your Jupyter notebook due April 28 (20% of the grade).