Group Proposal

1. What problem did you select and why did you select it?

   The goal is to create models that can predict the existence of an Exoplanet, using flux readings. It would be useful to compare our models to previous models done by others. It was also essential to use a different type of dataset rather than use images as done in previous classes and exams.

2. What database/dataset will you use? Is it large enough to train a deep network?

   NASA's Kepler Telescope (Time Series) Dataset is composed of a training and test set. The dataset is large enough to train a deep network as it consists of

   - 5087 rows or observations
   - 3198 columns or features
   - Column 1 is the label vector; Columns 2 - 3198 are the flux values over time
   - 37 confirmed exoplanet-stars and 5050 non-exoplanet-stars

3. What deep network will you use? Will it be a standard form of the network, or will you have to customize it?

   CNN, RNN(LSTM is a type of RNN)

   We plan to create at least three models, for each model we will customize the networks. And we will use the ensemble method if needed.

4. What framework will you use to implement the network? Why?

   Keras, Tensorflow

   Since Keras is a high-level neural network API, it is capable of running on top of TensorFlow. Both Keras and Tensorflow support convolutional networks and recurrent networks and the combination of these two as well.

5. What reference materials will you use to obtain sufficient background on applying the chosen network to the specific problem that you selected?

   Book 'NNDesign', lecture notes and other scholar articles.

6. How will you judge the performance of the network? What metrics will you use?

   Compare our model with those done by the others. The metrics would be accuracy, cohen kappa score, and F1 score of the validation set.

7. Provide a rough schedule for completing the project.

   First week: pick datasets, data preprocessing (Apr 5)

   Second week: data preprocessing, models(Apr 12)

   Third week: ensemble, other improvements(Apr 19)

   Forth week: report, slides(Apr 26)