Instructions for Study Participants

Summary

Using the data I've given you, build a machine learning model which has the best ROC AUC score you can achieve on the test set. Note: this is not a test of how well you can build machine learning models, it's an exercise in whether I can reproduce the model you create. Share your results in 3 stages: results and model architecture only, results plus model description, complete code.

Details

- I've already given you data and instructions on how to get the training, testing and validation data from the MIMIC-III dataset.
    - Please use this data, and don't do any pre-processing (reshaping is fine if necessary). The data has already been standardized and discretized. This ensures we're all using the same data.
- Build whatever model you want to, although please bear in mind that I will be trying to reproduce your model on a CPU. If you're using a GPU, let me know so I can report it in the abstract.
- Report whatever metrics you want to, but please also report ROC AUC (e.g. using sklearn.metrics.roc_auc_score) on the test set. This means I can compare all models using one consistent metric.
- When you have your completed model, send your results to me in 3 stages:
    - Give me your results and what architecture you used (but nothing else)
    - Give a description of your model (something like you'd find in an academic paper)
    - Send me a copy of your code
- These 3 stages can be in separate documents in the same email, whatever works for you, just as long as I can look at your results without knowing about your model and seeing your code.

I need results back by the 15th of March at the latest, so I have time to try and reproduce your model, do the analysis, write up the abstract and send to everyone for approval. Although if you can send me results back earlier I will be very much appreciative! Anyone who sends a model will get authorship.

Thanks!