Group Projects

Projects

<Option 1>:

- Pick a particular dataset of interest (demographics/age vs. drug use, mortality rates and causes in particular geographic regions, etc.)
    - Do feature selection, evaluate the data (demographics, distribution, features, etc.), consider 2-3 or more classification/regression techniques and apply them, evaluate the efficacy of each technique and compare results (i.e. sensitivity, AUC, feature ranking)

<Option 2>:

- Pick a ML/Data Science research area, find a paper that is interesting or in an area you want to learn in-depth
    - E.g. Bio-inspired Neural Networks, Prediction Uncertainty in Self-Driving Vehicles, Logistic Regression for Clinical Applications, X Supervised Learning Algorithm(s) on a particular task, etc.
- *Attempt* to follow the paper methodology to replicate the results of that paper
    - If you cannot replicate the results, show the results you did get, evaluate the data (data demographics, distribution, feature selection, etc.) and the models used (i.e. accuracy, ROC plots, feature ranking, etc.)

<Option 3>:

- Talk to me → if you're excited about something in ML/Data Science and want to do a project, I'm in!

OVERALL: the point is to pick a paper/dataset/topic, pre-process data, apply ML techniques, discuss and present

Projects

Paper submission (**due Tues, Dec 07 @ 11:59 pm EST**):

- Abstract, Introduction, Methodology, Results, References
- Size 12 font, single-spaced (2-column is welcome as well)
- No strict page length, but if you want a guideline, ~ 6 pages (not including references, appendix)

Presentations (**present on Thursday, Dec 02**):

- Format: Present your topic, methodology, results, etc,
- Thursday, Dec 02 starting @ 2:30 PM
- (5 mins per group + 2 min questions) * 15 groups ~ 105 minutes
- Set for in class, but may move to Zoom if we extend time per group

Sign up here (Groups of 4+):

- https://docs.google.com/spreadsheets/d/1wh5FWE-_7hLO5Sg8uGiXvCVvhS3-DzBx1Ay5VJ5Wvlo/edit?usp=sharing