CS 542 Final Project Ideas (2019)

Please look at the following for data sets:

1) UC Irvine data sets: http://archive.ics.uci.edu/ml/
2) Stanford Network data sets: http://snap.stanford.edu/data/index.html
3) Kaggle data sets: https://www.kaggle.com/datasets
4) Vast Data Sets: http://www.cs.umd.edu/hcil/varepository/benchmarks.php

Each of these the links above give you a data set to work on, list of papers that describe the research and ML tasks (classification, detection, clustering, etc.) people have done so far. Your project is to

a) read some of the list of papers to determine what the state of the art currently is on that data set.
b) Come up with a specific ML task you want to try on the data set
c) Design the ML algorithms to perform this task.

You will work on this project with a team of 3 and will submit the result by the day of the final:

- The report will be in the NIPS submission format which can be found in: http://nips.cc/Conferences/2013/PaperInformation/StyleFiles
- The code of your algorithms
- Power point presentation of slides between 5 to 10.

If you find the somewhat open ended research like problem above too hard, you can work on the more specific class projects from the following two course websites:

5) http://www.cs.cmu.exdu/~guestrin/Class/10701/projects.html
6) http://www.cs.cmu.edu/~ggordon/10601/projects.html

But I am hoping that most of you will choose a more research-y projects (1) – (4), as you may have the potential to turn it into a conference paper.

Please let our TA know what your team is going to work on by 3/5/19. However, You can start now! ☺ Please submit to our TA 's1 page summary of 1) your team (who is in your team), 2) problem you are going to work on, 3) a couple of rough ideas on your approach.