**Final Project Instructions**

**Final Project Ideas**

We should help each other with ideas for final projects. If you find something cool but end up not using it, suggest it to someone else. Post any helpful links or ideas on Google Classroom. Here are a few ideas off the topic of my head. If you happen to finish one topic early then you can start on a new one. The main things are 1.) keep yourself busy, 2.) keep yourself interested, and 3.) choose an appropriate level of challenge FOR YOU. Any topic you take on should be challenging but not impossible. You will also need to be relatively self sufficient as some of the topics below I don't even know much about myself!

1.) Use what you know:

- Analyze an interesting and complicated Kaggle dataset using the techniques that you have already learned.

2.) Learn more predictive techniques:

- Learn how to use neural networks (PyTorch, Keras, and TensorFlow are all options.)

- Learn other NLP applications such as chatbots

3.) Learn more technical stuff:

- Learn how to perform calculations on large datasets on the AWS cloud instead of on your local computer.

- Contribute to scikit-learn open source docs.

4.) Learn more technical and artsy stuff:

- Learn how to make more advanced D3 visualizations.

5.) Learn more theoretical stuff:

- Go into more of the theory behind the algorithms. This could involve understanding proofs and/or delving into more detail in multivariable calc, linear algebra, or statistics topics.

**Final Project Nuts & Bolts**

1. You need to keep an updated blog post for every day that we are in class describing a.) what you worked on that day, b.) any stumbling blocks you ran into, and c.) what you plan on working on next. This will both hold you accountable for managing your time well as long as inform me if you need my help.

2. You will need to communicate your final results through a.) well commented code, b.) a well written blog post, and c.) a clearly communicated oral presentation.

a.) The code should be extremely clear, well commented (with hashtags when necessary and markdown cells in between), and with no bugs. If you happen to be learning an entirely new skill such as neural networks, then think about modeling your notebook after the Jupyter Notebook lessons that I have provided for you this semester. (As in, reading your notebook will teach me a new skill).

b.) You need a final blog post that is as good as any medium or towardsdatascience article that we’ve read this semester explaining your work. This blog post should link to a GitHub code base repo. The blog will need to have images/charts/graphs in addition to text.

c.) You will need a short slide show to present during our assessment block. You will need to rehearse ahead of time, as we only have a limited amount of time for each presentation, and so your talk needs to take no more that the maximum allotted time.

3. This project will take 3 weeks and as such will be worth 200 project points.