Data Mining Project Milestones

Group 4 Members:
- Melissa Pollich
- Nhan Nguyen
- Benjamin Breyfogle

✓ **Milestone One: Data Collection**
- ❏ Description: find datasets with movie data
- *Complete by: July 12th, 2019*
- *Assigned to: Melissa, Nhan*

❏ **Milestone Two: Data Cleaning**
- ❏ Description:
  - ❏ Look at collected datasets and figure out important/relevant attributes
  - ❏ Build/organize new database with
    - ❏ Remove duplicates
    - ❏ Remove unnecessary data points
- *Complete by: July 19th, 2019*
- *Assigned to:*

❏ **Milestone Three: Data Analysis - Classification**
- ❏ Description:
  - ❏ Construct classification system based off of new cleaned/organized dataset
  - ❏ Divide dataset into *training dataset* and *test dataset*
- *Complete by: July 19th, 2019*
- *Assigned to:*

<mark>**PROJECT PART 3: PROGRESS REPORT - DUE JULY 26TH**</mark>

❏ **Milestone Four: Data Analysis/Prediction - Build Prediction Model**
- ❏ Description
  - ❏ Construct/Code prediction model using the training dataset to develop models which can be used for the test dataset
- *Complete by: July 26th, 2019*
- *Assigned to:*

- ❏ **Milestone Five: Data Analysis/Prediction - Test Prediction Model**
    - ❏ Description
        - ❏ Test prediction model code using test dataset for analysis using these prediction algorithms
            - ❏ Decision Tree
            - ❏ Regression
            - ❏ Neural networks
    - ● *Complete by: August 2nd, 2019*
    - ● *Assigned to:*

- ❏ **Milestone Six: Finalize Prediction Model**
    - ❏ Description
        - ❏ Finalize prediction model
        - ❏ Note changes
    - ● *Complete by: August 8th, 2019*
    - ● *Assigned to:*

==**PROJECT PART 4: FINAL REPORT, CODE & DESCRIPTIONS, PROJECT PRESENTATION - DUE AUGUST 9, 2019****==

- ❏ **Milestone Seven: Figures and Prediction Summary**
    - ❏ Description
        - ❏ Build graphs of the data from the prediction model test
        - ❏ Summarize results
    - ● *Complete by: August 9th, 2019*

Useful links

https://iopscience.iop.org/article/10.1088/1742-6596/1000/1/012100/pdf
https://www.kaggle.com/carolzhangdc/analyze-imdb-score-with-data-mining-algorithms#2-Data-Exploration