

Data Mining, CSCI 347, Fall 2019
Competitions Workshops, Oct. 18, Oct. 25 and Nov. 15

Updated 10/22/2019

Dalton and Hunter, Oct. 18

Competition: German Traffic Sign Recognition from a neural network research competition

Amanda, Keelie and Xaavan, Oct. xx

Competition: Classify Dirty Dishes from kaggle

Carson, Joseph and Kaleb, Oct. xx

Competition:

Working in pairs or triples, create a workshop to walk the class through a data mining completion. For the competition that you choose, one person give the url, describe the goal, ground rules, performance metric, they used, challenges, the format of the data, data characteristics, the format of submissions, and the successful approaches that were tried.

The other person then walk the class through the steps of one of the solutions.

If there is a 3rd person, they can walk the class through the steps of another solution.

These workshops will occur in a computer lab. All team members can walk around and help students complete the competition.

Here are some places to find data mining competitions:

- DrivenData: Data Science Competitions for Social Good <https://www.drivendata.org/>
- Kaggle, the leading platform for data prediction competitions <https://www.kaggle.com/>
- CrowdANALYTIX, converts business challenges into analytics competitions <https://www.crowdanalytix.com/>
- Innocentive, mainly focusing on life sciences, but has other interesting competitions <https://www.innocentive.com/>
- Tunedit, education, research and industrial contests. <http://tunedit.org/>

This will be graded as follows:

The point breakdown will be:

<p>Overview (30 points)</p> <ul style="list-style-type: none">UrlGoalGround rulesPerformance metricChallengesFormat of the dataData characteristicsFormat of submissionsSuccessful approachesVisual aids
<p>Walk through (70 points)</p> <ul style="list-style-type: none">Advance setup (if needed)Handout for walk throughClarity of stepsTimingCompletenessLearning