UMASS AMHERST

Completed • Knowledge • 41 teams

UMass Twitter NER

Thu 6 Nov 2014 – Mon 15 Dec 2014 (8 months ago)

Dashboard

Home
Data
Make a submission

Information
Description
Evaluation
Rules

Forum

Leaderboard

My Team

My Submissions

Leaderboard

1. Skeuomorph
2. Aysha
3. @JustinBieber
4. Schattie
5. pighead
6. NGU
7. CrazyBird
8. Aha
9. catscatsBORSCHTcats
10. ominousmango

Forum (0 topics)

41 teams

49 players

104 entries

This competition is private-entry. You've been invited to participate.

Evaluation

The evaluation metric for this competition is Mean F1-Score. The F1 score, commonly used in information retrieval, measures accuracy using the statistics precision p and recall r. Precision is the ratio of true positives (tp) to all predicted positives (tp + fp). Recall is the ratio of true positives to all actual positives (tp + fn). The F1 score is given by:

$$F1 = 2\frac{p \cdot r}{p + r} \quad \text{where} \quad p = \frac{tp}{tp + fp}, \quad r = \frac{tp}{tp + fn}$$

The F1 metric weights recall and precision equally, and a good retrieval algorithm will maximize both precision and recall simultaneously. Thus, moderately good performance on both will be favored over extremely good performance on one and poor performance on the other.

For NER, evaluating F1 is better than evaluating token accuracy, since the vast majority of tokens are not in a named entity.

Submission Format

See the documentation on the course webpage. We will provide python docs for converting the output of crfsuite to what kaggle expects.