"everything you 'd expect -- but nothing more"

Completed • Knowledge • 861 teams

Sentiment Analysis on Movie Reviews

Fri 28 Feb 2014 – Sat 28 Feb 2015 (4 months ago)

Dashboard

Home
- Data
- Make a submission

Information
- Description
- Evaluation
- Rules

Forum

Leaderboard

My Team
- GitHub

My Submissions

Leaderboard

1. Mark Archer
2. Armineh Nourbakhsh
3. Merlion
4. Puneet Singh
5. Yoon
6. DrStrangelove
7. akqwerty
8. MDAKMlab
9. st_sopov
10. JR

Forum (41 topics)

Split tweet into phrases
55 days ago

Stanford NLP - Training is not converging.
55 days ago

Release test dataset labels ?
2 months ago

Our text summarizer and keyword generator is available for use
3 months ago

So how did you implemented it ?
4 months ago

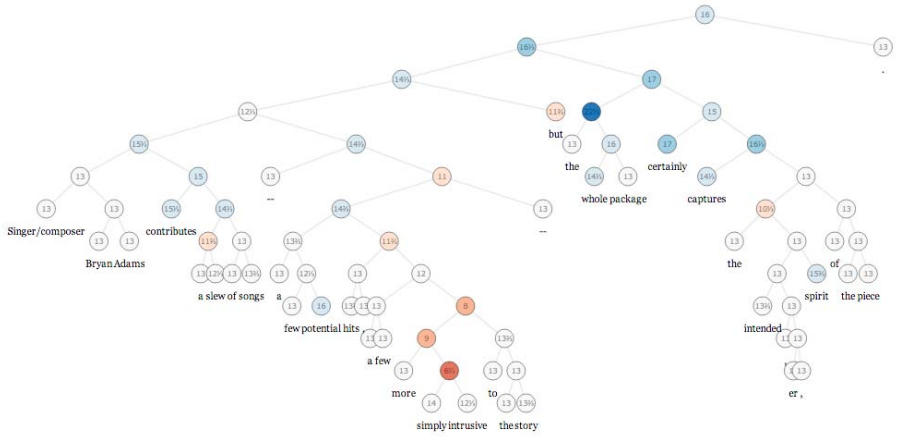
Looking for a partner to build a recurrent neural network

Competition Details » Get the Data » Make a submission

Classify the sentiment of sentences from the Rotten Tomatoes dataset

"There's a thin line between likably old-fashioned and fuddy-duddy, and The Count of Monte Cristo ... never quite settles on either side."

The Rotten Tomatoes movie review dataset is a corpus of movie reviews used for sentiment analysis, originally collected by Pang and Lee [1]. In their work on sentiment treebanks, Socher et al. [2] used Amazon's Mechanical Turk to create fine-grained labels for all parsed phrases in the corpus. This competition presents a chance to benchmark your sentiment-analysis ideas on the Rotten Tomatoes dataset. You are asked to label phrases on a scale of five values: negative, somewhat negative, neutral, somewhat positive, positive. Obstacles like sentence negation, sarcasm, terseness, language ambiguity, and many others make this task very challenging.



Kaggle is hosting this competition for the machine learning community to use for fun and practice. This competition was inspired by the work of Socher et al [2]. We encourage participants to explore the accompanying (and dare we say, fantastic) website that accompanies the paper:

http://nlp.stanford.edu/sentiment/

There you will find have source code, a live demo, and even an online interface to help train the model.

[1] Pang and L. Lee. 2005. *Seeing stars: Exploiting class relationships for sentiment*

4 months ago

teams

players

entries

categorization with respect to rating scales. In ACL, pages 115–124.

[2] *Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank*, Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Chris Manning, Andrew Ng and Chris Potts. Conference on Empirical Methods in Natural Language Processing (EMNLP 2013).

Image credits: Popcorn - Maura Teague, http://www.flickr.com/photos/93496438@N06/

Started: 4:54 pm, Friday 28 February 2014 UTC
Ended: 11:59 pm, Saturday 28 February 2015 UTC (365 total days)
Points: this competition did not award ranking points
Tiers: this competition did not count towards tiers