Antimony or Antinomy?

NLP Classification of Science and
Philosophy Subreddits
Mac McCarthy

About

Scraping the web

Preparing the data

Building the model

Testing the model

Reddit Data

-Discussion forum: formal versus informal

-API

r/science versus r/philosophy

-Names?

-Words?

-Phrases?

Preparing the data

Cleaning up words

-Common stopwords

-"Science" and "philosophy"

-No stemming or lemmatizing

Splitting

-1500 train, 500 test

Vectorizing

-Counting

-TF/IDF

Models

- Logistic Regression

- K-Nearest Neighbors

- Random Forests

- Extra Trees

- Bagging

- Adaptive Boosting

- Support Vector Machines

- Multinomial Bayes

Solution

-Logistic Regression: 99.4%

-K-Nearest Neighbors: 99.0%

-Random Forests: 99.4%

-Extra Trees: 99.4%

-Bagging: 99.4%

-Adaptive Boosting: 99.6%

-Support Vector Machines: 99.4%

-Multinomial Bayes: 99.2%

Adaptive Boosting!

What's important?

	importances	features
670	0.085714	first
1665	0.057143	study
185	0.057143	bone
1832	0.057143	used
720	0.057143	future
1526	0.057143	scientists
565	0.057143	every
1571	0.042857	show
143	0.042857	bacteria
1615	0.028571	species

Fun stuff

Donald J. Trump ✔
@realDonaldTrump

Follow

The Woodward book is a scam. I don't talk the way I am quoted. If I did I would not have been elected President. These quotes were made up. The author uses every trick in the book to demean and belittle. I wish the people could see the real facts - and our country is doing GREAT!

4:32 AM - 7 Sep 2018

Fun stuff

Cher: So like, right now for example. The Haitians need to come to America. But some people are all, "What about the strain on our resources?" Well it's like when I had this garden party for my father's birthday. I put R.S.V.P. 'cause it was a sit-down dinner. But some people came that like did not R.S.V.P. I was totally buggin'. I had to haul ass to the kitchen, redistribute the food, and squish in extra place settings. But by the end of the day it was, like, the more the merrier. And so if the government could just get to the kitchen and rearrange some things we could certainly party with the Haitians. And in conclusion may I please remind you it does not say R.S.V.P. on the Statue of Liberty! Thank you very much.

Questions

Thanks!