

Social Media Sentiment Analysis: Detection of Negative/Depressive Messages on Twitter

—

Malik and Diego

REAL-WORLD PROBLEM

- . Is it possible to identify Tweets with a negative or depressive connotation using Machine Learning?
- . If so, this project can aid researchers to utilize social media postings as a method to identify individuals with trouble and apply early interventions.

The Experiment

- We will use a data set from Kaggle.com that contains over 1.6 million tweets annotated on a scale from 0 (negative) to 4 (positive). 1 million for training and the rest is test data
- We will train the model to recognize what words have a negative/depressive connotation and which are positive
- We will use algorithms such as Naive Bayes and Logistic —Regression to rank tweets on a positive to negative scale
- After that we will try to improve our model to be more accurate by taking into account word context and other factors

Materials

Python Libraries

- Python
- Pandas
- Matplotlib
- Seaborn
- Numpy
- Sklearn
- Logistic Regression
- NLTK
- Gensim
- Word2Vec

Hypothesis

We hypothesize that we will be able to accurately predict the connotation of Tweets using

— Machine Learning.

FEASIBLE?

- . The students embarking on this project have intermediate coding skills, which should be enough to produce a successful and insightful project outcome.
- . One of the group members is proficient at sentiment analysis.
- .The other group member will get up to speed by taking short online courses on sentiment analysis prior to the analysis.
- .The professor advising this project determined that the project is feasible.

References (Primarily websites)

1. <https://www.nltk.org/howto/sentiment.html>
2. <https://monkeylearn.com/sentiment-analysis/>
3. <https://www.kaggle.com/kazanova/sentiment140/kernels>
4. <https://time.com/1915/how-twitter-knows-when-youre-depressed/>
5. <https://www.psychiatry.org/psychiatrists/practice/dsm>
6. https://en.wikipedia.org/wiki/Sentiment_analysis
7. <https://www.datacamp.com/community/tutorials/simplifying-sentiment-analysis-python>
8. https://link.springer.com/chapter/10.1007/978-3-642-40319-4_18
9. https://link.springer.com/chapter/10.1007/978-3-319-49586-6_59
10. <https://www.datacamp.com/courses/analyzing-social-media-data-in-python>