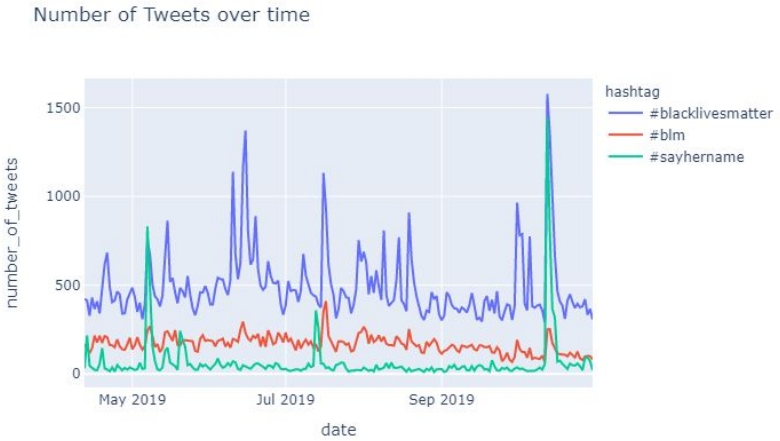
If You're Happy and I Know it

Sentiment Analysis of Tweets under Black Lives Matter (BLM) Hashtags

Peggy (Zijin) Wang '21
Data Science Capstone
CW: Police Brutality

Background

- Research project: studying tweets across 37 BLM hashtags
    - key dates and protests
- What are people tweeting about and how are they using different hashtags?
- **What is the distribution of sentiments across BLM hashtags?**
    - Sentiments: positive, negative, neutral



May 2019- Video of Sandra Bland's arrest released
October 13, 2019- Murder of Atatiana Jefferson

Data

- Collected with Tweepy, Twitter API
- Tweets from #blm, #blacklivesmatter, and #sayhername
- 2019-04-12 to 2019-10-31 (30 weeks)
- n=146 682

#blacklivesmatter	99792
#blm	33912
#sayhername	12978

- Only kept tweets with 3 words or more
    - Super short tweets often only contained the hashtag + an external link

Sentiment Analysis Models

Supervised Learning (Classification):

- Naive Bayes
- Logistic Regression

Training Data:

- Kaggle dataset
- 1.6 million labeled tweets
  - perfect split between + and -
- Randomly sampled 160 000 tweets
  - model accuracy was the same

Existing Sentiment Library:

- VADER (Valence Aware Dictionary and sEntiment Reasoner)

Feature Extraction: Bag of Words

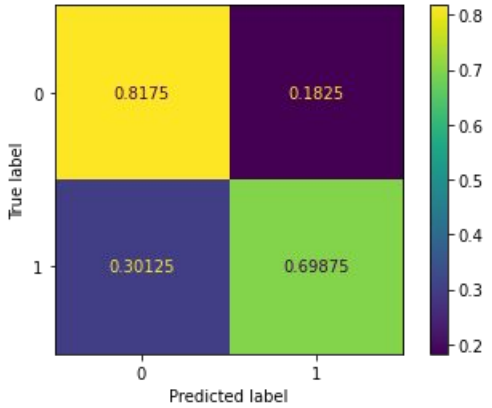
- Dictionary represented as a matrix. Each column is one word
- Example:
    - 1. “I love Data Science”
    - 2. “I love Math”
- 136501 features (columns)

I	love	Data	Science	Math
1	1	1	1	0
1	1	0	0	1

0 = negative sentiment
1 = positive sentiment

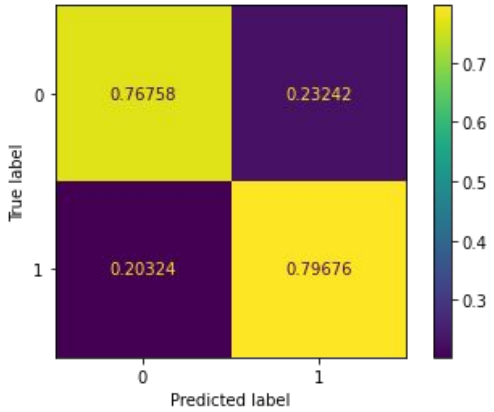
Training

Naive Bayes Classifier



Accuracy: 0.76718 (10-fold cv)

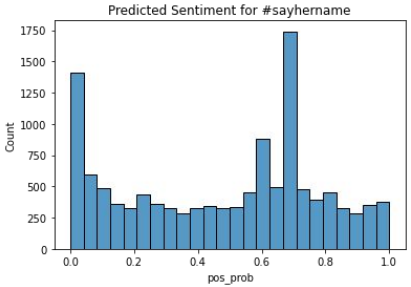
Logistic Regression

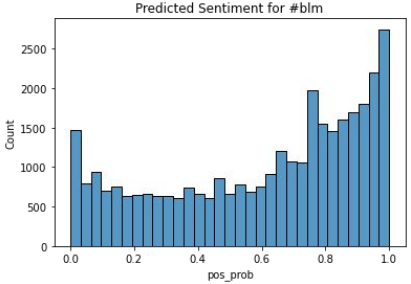


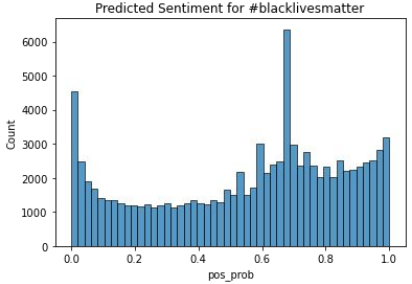
Accuracy: 0.78226 (10-fold cv)

Predictions

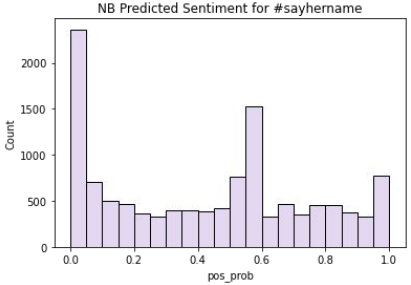
Logistic Regression:

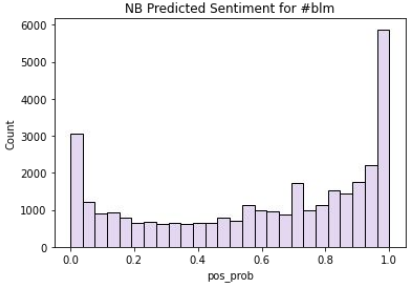


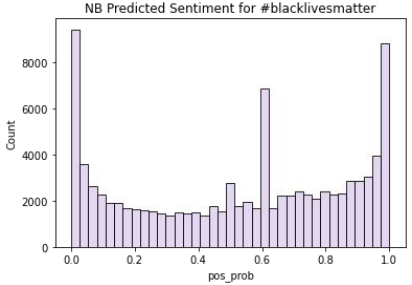




Naive Bayes:







Evaluation

- Manually labeled data, 100 tweets from each hashtag
- High rate of false positives (~0.5) across both models
- Difficult to deal with neutral sentiment
- Tried labeling tweet as neutral if P(class 1) is closest to 0.5 but accuracy was still very low:

Logistic Regression Performance

```
#sayhername 0.54
#blm 0.35
#blacklivesmatter 0.39
```

Naive Bayes Performance

```
#sayhername 0.47
#blm 0.31
#blacklivesmatter 0.38
```

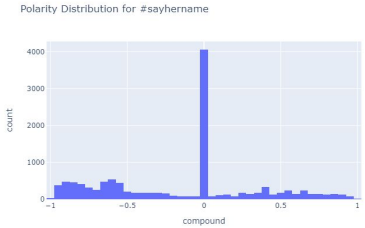
Using Existing Packages

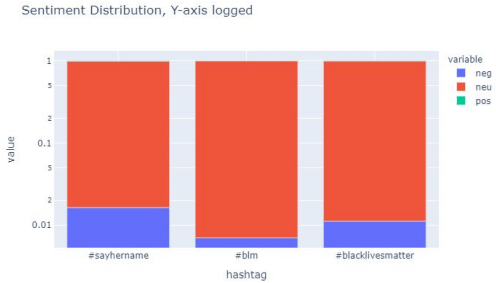
VADER (Valence Aware Dictionary and sEntiment Reasoner)

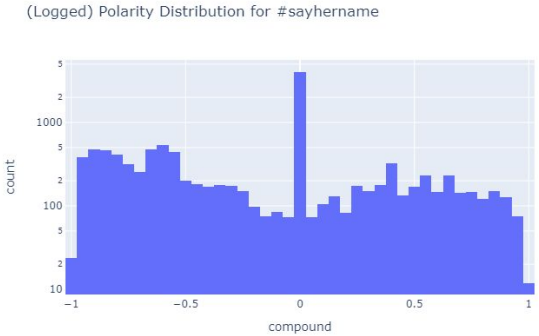
- Specifically for social media analysis
- Returns a compound score from -1 to 1 denoting sentiment
- Takes into account emojis, negation, punctuation, etc.

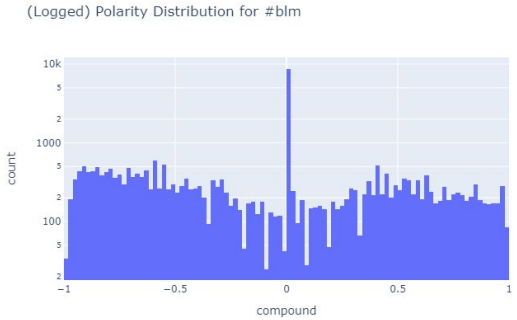
Predictions I

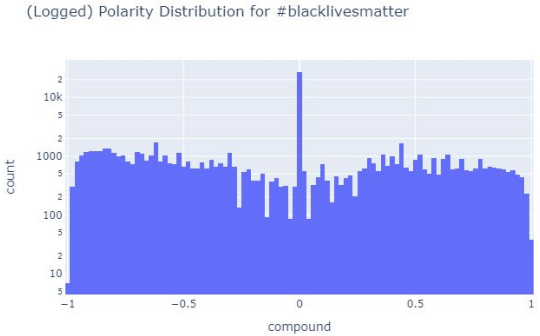
Unlogged:





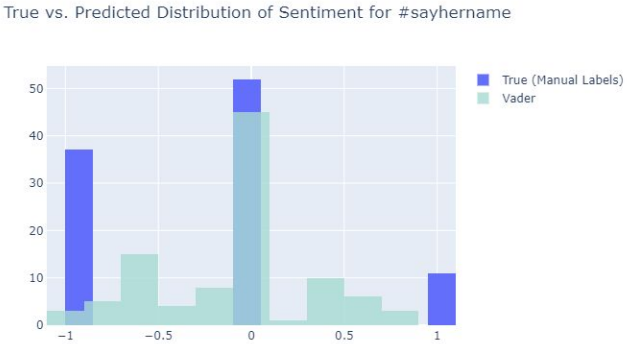


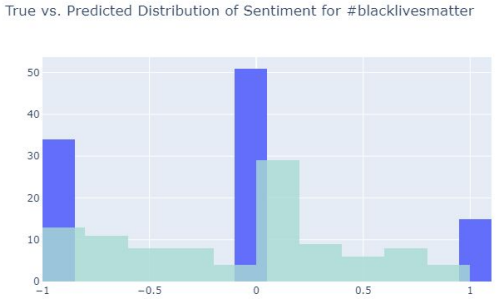




Predictions II

- Compared my labels to VADER output
- Not always easy to label
  - Vader giving more conservative classifications

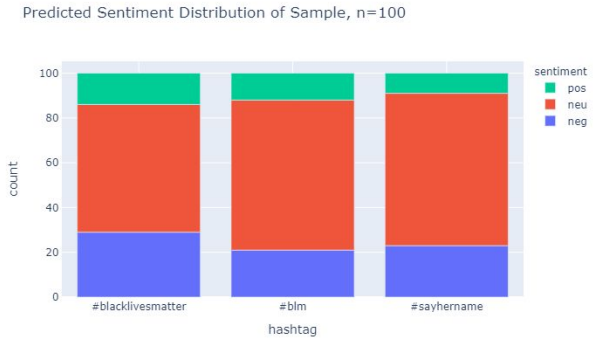




Evaluation

Hashtag	Labeled	Predicted (Vader)
#sayhername	Negative: 37 Neutral: 52 Positive: 11	23 68 9
#blm	Negative: 34 Neutral: 60 Positive: 6	21 67 12
#blacklivesmatter	Negative: 35 Neutral: 51 Positive: 15	29 57 14

Accuracy
Average: 0.62667
#sayhername: 0.63
#blm: 0.64
#blacklivesmatter: 0.61



Results

- According to VADER, mostly neutral tweets across all 3 hashtags
    - More negative sentiment than positive reflecting seriousness of topic

- Difference between classifiers and VADER: Nuances in language
    - Classifiers only looking at the corpus

Discussion + Limitations

- Difficult to label sentiment → personal biases
- Some tweets have both positive and negative sentiment, others use sarcasm
- Low model accuracy
- Future work: Multi-class classification (deep learning), Topic Modeling

Thank You for Listening!