

Sentiment Analysis | Business Reviews

Miguel Angel Santana II, MBA

February 2, 2021

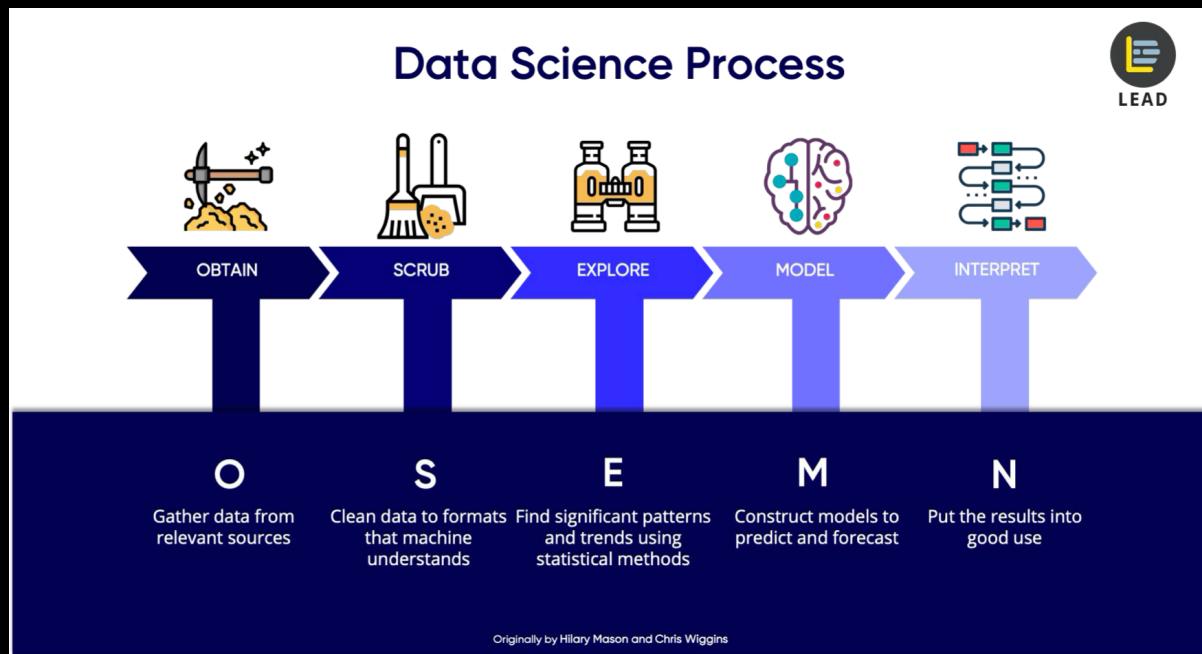
Introduction

- Methodology
- OSEMN Framework
 - Obtain | Scrub
 - Exploratory Data Analysis
 - Sentiment Modeling
 - Interpret Results
- Limitations & Future Work
- Thank You

Methodology

- Leverage OSEMN to process, model and gain insight into NLP practices.
- Process Text | Sentiment Analysis
 - Amazon Reviews
 - Movie Reviews
- Compare Predictions & Pre-determined Labels
- Compare Text Classifier and Vader Sentiment Analyzer

OSEMNN Framework



Obtain | Scrub

Stanford Artificial Intelligence Laboratory
(SAIL)

- <http://ai.stanford.edu/~amaas/data/sentiment/>

Udemy Course: NLP – Natural Language Processing

- by Jose Portilla.

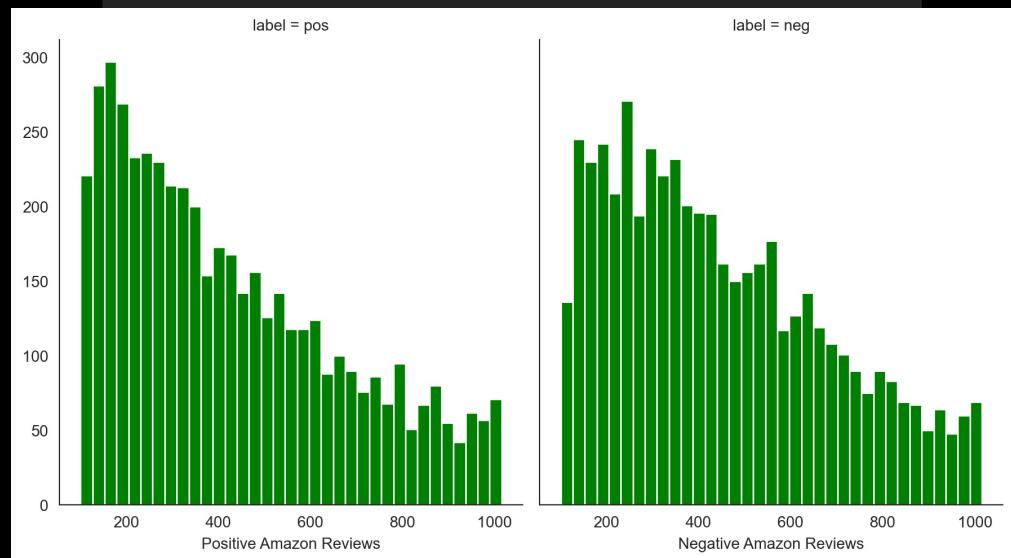
Scrubbing | Data Cleaning

- Key Decisions
 - drop whitespace strings
 - drop null values

Exploratory Data Analysis

Amazon Product Reviews

- Positive and negative reviews illustrate a similar length in text.

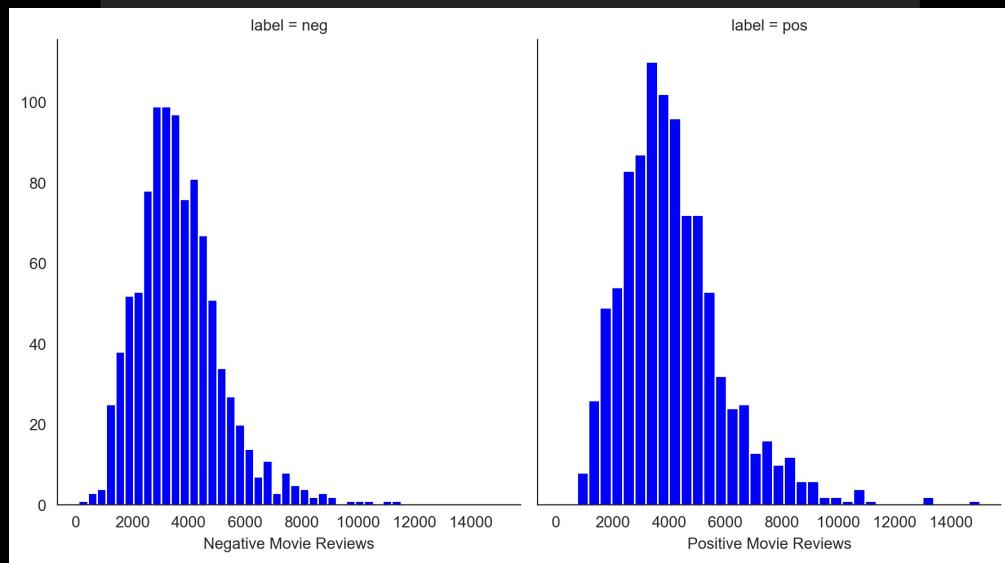


Exploratory Data Analysis

Continued

Length of Movie Reviews

- Positive and negative movie reviews have a similar distribution with positive reviews reflecting more outliers.



Sentiment Modeling | Vader Analyzer

Valence Aware Dictionary for Sentiment Reasoning is a model used for text sentiment analysis in natural language processing. The model offers insight into polarity as well as intensity

Amazon Reviews

	precision	recall	f1-score	support
neg	0.86	0.51	0.64	5097
pos	0.64	0.91	0.75	4903
accuracy			0.71	10000
macro avg	0.75	0.71	0.70	10000
weighted avg	0.75	0.71	0.70	10000

The overall accuracy score is 0.7091.

Movie Reviews

	precision	recall	f1-score	support
neg	0.72	0.44	0.55	969
pos	0.60	0.83	0.70	969
accuracy			0.64	1938
macro avg	0.66	0.64	0.62	1938
weighted avg	0.66	0.64	0.62	1938

The overall accuracy score is 0.6367389060887513.

Interpret Results

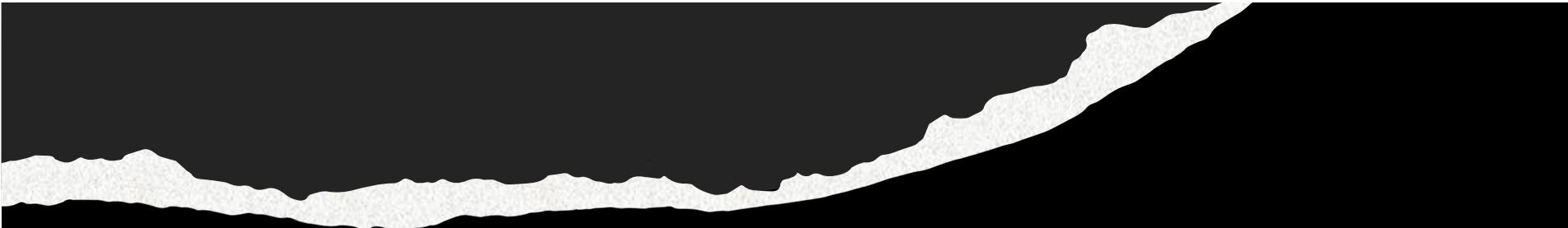
The Vader sentiment analyzer was unable to outperform the text classifier in Github project “Text Classification – Movie Reviews” with classified text with a 92% accuracy rate.

The analysis of the amazon product reviews was more successful than the movie reviews.

Movie reviews typically have qualities that make Vader analysis challenging.

Limitations and Future Work

- Vader sentiment analysis struggles with positive and negative sentiment within the same text.
 - Movie reviews are prone to positive and negative sentiment due to differences in quality of acting versus quality of movie script.
- Vader struggles with sarcasm.
 - Using positive words in a negative way.
- Future work should include sentiment analysis in combination with speech-to-text and text classification.



Thank You!

Questions? Miguel Santana | Contact: santana2.miguel@gmail.com

Additional projects can be found on Github.

Username: [miguelangelsantana](#)