

Semantics and Sentiment Analysis

Section Goals

- Understand Sentiment Analysis
- Leverage Sentiment Analysis for Text Classification

Unlabeled Text Sentiment

- We've already explored text classification and using it to predict sentiment labels on pre-labeled movie reviews.
- But what if we don't already have those labels ?
- Are there methods of attempting to discern sentiment on raw unlabeled text?

VADER

(Valence Aware Dictionary for sEntiment Reasoning)

Sentiment with Python and NLTK

Vader for Sentiment Analysis

- VADER is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion.
- It is available in the NLTK package and can be applied directly to unlabeled text data.

VADER

- Primarily, VADER sentiment analysis relies on a dictionary which maps lexical features to emotion intensities called sentiment scores.
- The sentiment score of a text can be obtained by summing up the intensity of each word in the text.

VADER understand the strength of Emotion

- For example, words like “love”, “like”, “enjoy”, “happy” all convey a **positive** sentiment.
- VADER is intelligent enough to understand basic context of these words, such as “**did not love**” as a negative sentiment.
- It also understands capitalization and punctuation, such as “**LOVE!!!!**”

Sentiment Analysis Challenge

- Sentiment Analysis on raw text is always challenging however, due to a variety of possible factors:
 - Positive and Negative sentiment in the same text data.
 - Sarcasm using positive words in a negative way.

Let's explore using VADER sentiment analysis
with NLTK and Python!