Model Performance

Model Name: Vader **Test Date:** 23/03/2022 15:59:30 **Creator:** Tobias Rothlin

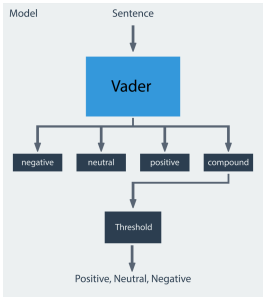
Overview

ML Principle:
Vader Sentiment Analysis

References:

- Vader overview
- Vader github

Algorithm Description:
Vader (Valence Aware Dictionary for sEntiment Reasoning) is a pre trained model used for sentiment analysis. Vader is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. The backbone of Vader is a dictionary that maps lexical features to emotion intensities (sentiment score). To receive the sentiment score of a sentence the intensities of each word are added. For example, words like 'love', 'enjoy' indicating a positive sentiment. Vader is smart enough to understand basic context like 'did not love' as negative. Further it has a basic understanding of capitalization and punctuation to emphasis tone. Due to this any preprocessing steps should not be done.



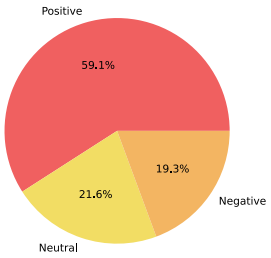
Classification Pipeline

Metrics

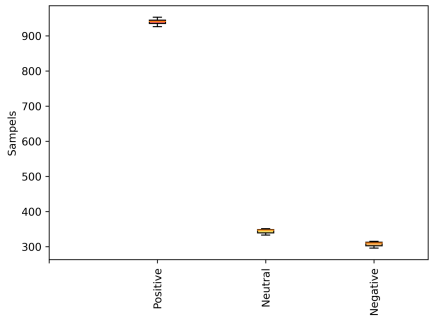
Data: ClassifiedDataSetV1.2 with 10 folds cross validation
Split seed: 4.83819

Training Dataset
(average)

Classes	Number of samples
Positive	940
Neutral	343
Negative	307



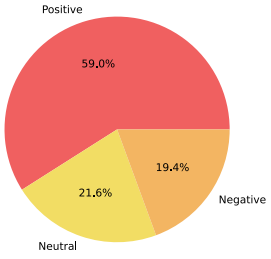
Average distribution of the samples



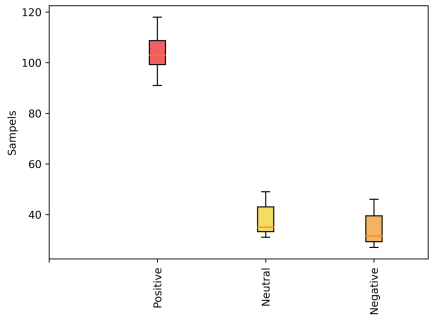
Distribution of the samples contained in each test split

Test Dataset
(average)

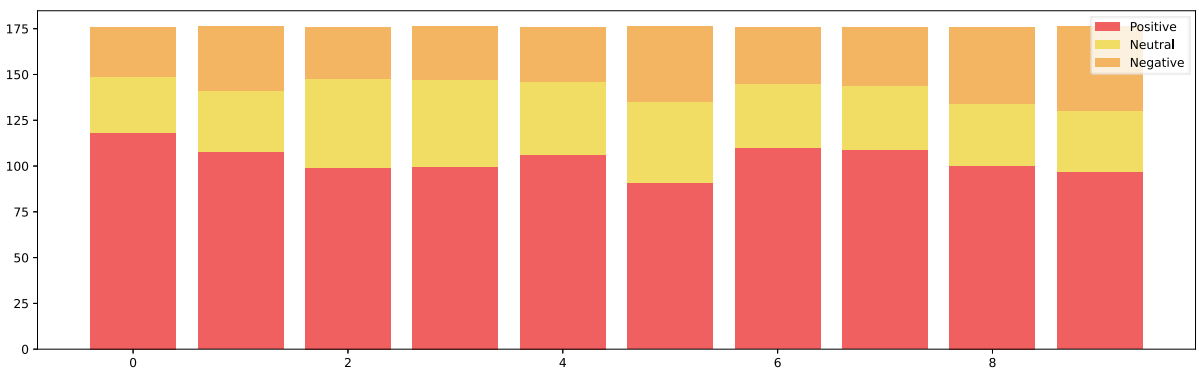
Classes	Number of samples
Positive	103
Neutral	38
Negative	34



Average distribution of the samples



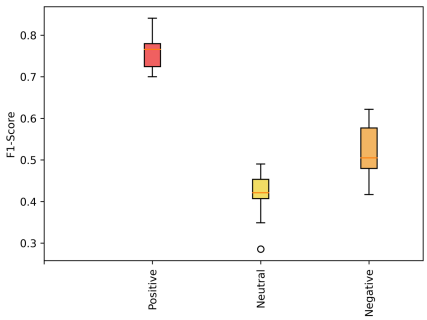
Distribution of the samples contained in each test split



Detailed training split composition

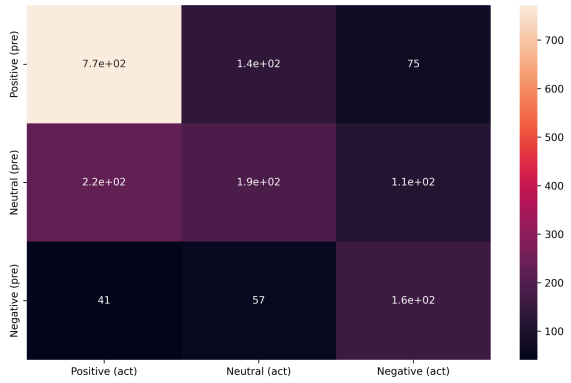
Classification Performance

Classes	Precision	Recall	F1 Score
Positive	78.62%	74.37%	76.44%
Neutral	36.07%	49.61%	41.77%
Negative	61.42%	45.75%	52.44%
Accuracy			63.47%
Macro Average	58.70%	56.58%	56.88%
Weighted Average	66.07%	63.47%	64.28%

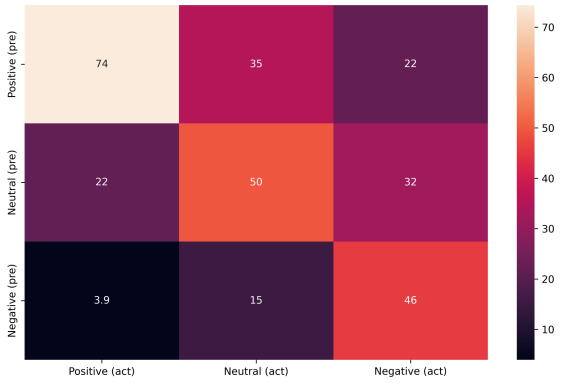


Distribution of the F1-Score

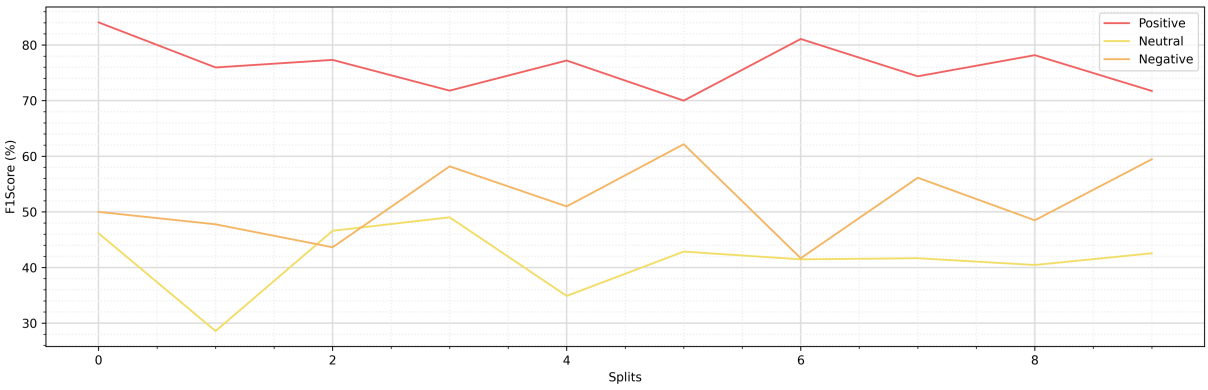
ConfusionMatrix:



Normalised ConfusionMatrix:



F1 Socre by split:



F1-Score per split