

Sentimental Analysis of Amazon Reviews

Sagar Punn
Computer Science
Ryerson University
Toronto, Canada
sagar.punn@ryerson.ca

Adam Sorrenti
Computer Science
Ryerson University
Toronto, Canada
adam.sorrenti@ryerson.ca

Saleh Al-Sailani
Computer Science
Ryerson University
Toronto, Canada
salsailani@ryerson.ca

Eisa Keramatinejad
Computer Science
Ryerson University
Toronto, Canada
eisa.keramati@ryerson.ca

Abstract— Proposal of an application level project using machine learning classification algorithms in sentimental analysis and comparing results between algorithms.

I. INTRODUCTION

Sentiment analysis is the interpretation and classification of text based data. The point of this analysis is to categorize each datapoint into a class that represents it's quality (positive, negative, etc.). Sentiment analysis can focus on polarity, emotions and intentions of customers. Classic Sentiment Analysis consists of the following steps: preprocessing, training, feature extraction and classification. The method used in this paper will follow the classical approach.

a) Prior Work: Reference [1], a comparative study, implemented Support Vector Machines (SVM), Naive Bayes (NB) and Maximum Entropy (ME) for classification of online reviews. It will be interesting to compare our the efficacy algoirhtm verses in appraoch taken in the article mentioned above.

II. MOTIVATION

Putting machine learning algorithms to use in order to predict the sentiment of human feedback is an interesting and emerging research area that has many applications. We chose to predict the sentimental value of Amazon reviews [2] based on the frequency of good/bad words and the length of the review using the one to five star user inputted value as our label. This will be an application level project that applies machine learning to sentimental analysis and performs comparative analysis between various algorithms. Presenting the efficacy of machine learning algorithms as they apply to sentimental analysis is a key motivation.

This problem is important as it can be applicable to industry in many ways such as marketing return on investment analysis. Understanding customers' feelings about a certain product is essential to business success and nowadays with the popularity of online reviews, companies are looking into convenient ways to process and analyze these reviews.

III. METHOD

Initial approach for feature extraction will be counting the total number of positive and negative words in a comment, because positive and negative words frequency (with reference to positive/negative words lexicons[3]) could have a high part in

dictating the sentiment. In addition to these, length of a review is also being considered one of the features. This will be our initial approach to the experimentation with plans to extract complex features.

We are then (after initial experiment) planning to use the bag of words model for our features. For our classification and training algorithms we are thinking between Probabilistic models like naive bayes, K Nearest Neighbors clustering algorithm and support vector model, we might add to or change these algorithms later on in the project.

- 1) For the preprocessing part of our algorithm we start with the basic text preprocess. This step consists of tokenization, normalization and noise removal.
- 2) We are hoping that by the time we start this part of the project we have learnt some basics of feature extraction so that we can apply it to our own project.
- 3) For training and classification we decided to use the above algorithms and compare the final results.

Using the bag of words model allows us to treat each of the terms in the vocabulary as a feature. Reviews will be assigned a value of 0 to with a rating below 3 stars (indicating that the product is in the bad category), or a value of 1 to reviews with ratings above 3 stars (meaning this product belongs to the good category). For each review with 3 stars we are going to randomly assign it a value of either 0 or 1. This way we will have our labels for both the training and evaluation.

IV. INTENDED EXPERIMENTS

This project aims to experiment with multiple classification algorithms such as, logistic regression, SVM, KNN. The results will then be compared to find out why each algorithm scored such accuracy. Additionally, the aim is to then improve on preprocessing to extract more features which can then improve the accuracy of the results compared to previous experimentation. The dataset will be divided into testing and training sets (e.g. 70:30 % split), with training data being fed into each algorithm. Afterwards, testing data will be used to make predictions which will be evaluated with an accuracy score (%) while also utilising a variety of loss functions. We will use the actual labels of the test dataset and compare it to the corresponding label provided by the algorithm for each point in the dataset to find the accuracy score. The raw dataset takes the

form of {"overall","reviewText","summary"} with adequate preprocessing we intend on train algorithms using the following structure {"positive word frequency", "negative word frequency", "reviewTextLength", "rating"}.

V. PLANNING AND MILESTONES

A. Planning

The project will be done by four members randomly assigned in the class of CPS803 at Ryerson University. Tasks will be done more collaboratively than individually. As we're in a social distanced environment, all meetings will be done virtually. Voice channels in Discord would be the main area of communication.

B. Milestones

1) *Milestone 1:* (Oct 10): [Preprocessing] Since the project deals with a lot of text, figuring out the best way to preprocess the data from the dataset (Sagar and Saleh).

2) *Milestone 2:* (Oct 20): [Feature Extraction] try to find and incorporate more complex features if possible (Adam and Eisa).

3) *Milestone 3:* (Oct 30) [Trial] successful application of one algorithm (Naive Bayes) with the preprocessed dataset and

features to get the accuracy of the predictions with relation to the labels (all members).

4) *Milestone 4:* (Nov 15) Evaluating and applying all different machine learning algorithms we could use (all members).

5) *Milestone 5:* (Nov 30) Comparatively analyse the predicted results of the different applicable machine learning algorithms and failure analysis of failed algorithms (all members).

6) *Milestone 6:* (Dec 15) Completing the report, 5-minute video, and poster if time permitted (all members).

REFERENCES

- [1] Rathor, A. S., Agarwal, A., & Dimri, P. "Comparative Study of Machine Learning Approaches for Amazon Reviews." 2018
- [2] Jianmo Ni, Jiacheng Li, Julian McAuley. "Justifying recommendations using distantly-labeled reviews and fine-grained aspects." Empirical Methods in Natural Language Processing (EMNLP), 2019.
- [3] Bing Liu. "Opinion Mining." Invited contribution to Encyclopedia of Database Systems, 2008.