SENTIMENT ANALYSIS OF ONLINE REVIEWS FROM YELP OPEN DATASET

By

Yu Heng Wu, BSc, University of Toronto Scarborough 2021

A Major Research Project

Presented to Toronto Metropolitan University (formerly Ryerson University)

In partial fulfillment of the requirements for the degree of

Master of Science

In the Program of

Data Science And Analytics

Toronto, Ontario, Canada, 2023

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A MAJOR RESEARCH PROJECT (MRP)

I hereby declare that I am the sole author of this Major Research Paper. This is a true copy of the MRP, including any required final revisions.

I authorize Toronto Metropolitan University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Toronto Metropolitan University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

Yu Heng Wu

Abstract

In the digital age, online reviews have become a central component in driving consumer choices. This study focuses on sentiment analysis of Yelp reviews, juxtaposing traditional machine learning (ML) algorithms (Naïve Bayes, Logistic Regression, Random Forest, Support Vector Machines) against the contemporary BERT model. Drawing from a vast dataset of over 6 million reviews, a balanced training set was derived by undersampling prevalent 5-star reviews. Our key objectives encompass both categorizing reviews into positive or negative sentiments, but also predicting precise star ratings. Remarkably, while conventional ML models demonstrated a range of accuracy levels, BERT stood out with its proficiency, particularly in positive/negative sentiment classification, reaching a flawless accuracy rate. These findings underscore BERT's potential in complex sentiment tasks, even as traditional models showcase notable abilities. The performance of each model is evaluated based on classification reports and a confusion matrix.

Keywords: Sentiment Analysis, Natural Language Processing(NLP), Review Classification, Comparative Analysis

Introduction

A. Background

In the digital age, online reviews have become the go-to medium for consumers to voice their opinions and share experiences, influencing the decisions of countless other consumers. Platforms like Yelp have transformed the way businesses receive feedback, offering instantaneous insights into the strengths and shortcomings of products and services. Given the vast volume of this user-generated content, it becomes clear that there is a need to deploy advanced computational techniques to process these reviews in order to filter the noise and capture consumer sentiments.

B. Research Question

This project seeks to conduct Sentiment Analysis on the Yelp Open Dataset. By comparing the classification reports and confusion matrices, we seek to determine the efficacy and accuracy of several state-of-the-art algorithms for discerning sentiment from raw text. The algorithms selected for this project are Naïve Bayes, Logistic Regression, Random Forest, Support Vector Machines, and the deep learning model BERT. This study will conduct the binary classification of sentiments into positive and negative, as well as utilizing the above algorithms in predicting precise star ratings, ranging from 1 to 5 stars. This project was written and executed in Python, the specifics of the code can be found in the GitHub link found in Appendix A.

Literature Review

This project seeks to leverage the power of machine learning to analyze and compare the sentiment of restaurant reviews using various machine learning and deep learning techniques. This literature review aims to review articles related to this MRP in the domain of sentiment analysis (SA), articles used as references for data exploration of the dataset, and articles on some of the Machine Learning and Deep Learning techniques that are used in this MRP.

I began by searching for papers and articles on SA conducted on online user review data. I began looking for SA of text reviews that contain a star rating since for this project, I will be training the sentiment classifiers by utilizing the star ratings on the Yelp reviews as a label for the sentiment of the review text. I began with “Thumbs up? Sentiment classification using machine learning techniques” [1] by Pang et al. which utilized a set of movie reviews containing star ratings. They manually chose words to determine the sentiment classification for the text data of those reviews and compared those results to the results of Naive Bayes, Maximum entropy, and SVM algorithms. For this project I wanted to use the star ratings as a label for the sentiment of the text review, similar to a paper done by Alec Go et al. in “Twitter Sentiment Classification using Distant Supervision”[4] they trained their sentiment classifiers using 1.6 million tweets with emoticons within the tweets as the label for their sentiment. In another article, “Predicting the Sentiment Polarity and Rating of Yelp Reviews” [2] Andrew Elkouri utilized 3 machine learning models (Naive Bayes, Support Vector Machines, and Logistic Regression). He trained two classifiers using the star ratings as their labels, one classifier to perform positive/negative classification and the other to predict 5-star classification. Both of these papers utilized distant supervision for their sentiment classifiers. I also plan to use a mix of Machine learning and Deep learning techniques, similar to “Sentiment Analysis of Twitter Data Using TF-IDF and Machine Learning Techniques” [12] by S.Singh et al, where SA of Twitter data was done the various algorithms were then compared to each other based on their accuracy.

Data exploration

Before doing any analysis or classification, it is critical to conduct data exploration. An in-depth exploration of the data offers a chance to familiarize myself with my data, identify potential issues, feature engineering, and assists in deciding which analysis techniques to use. A detailed exploratory analysis of the dataset was done using references from “Descriptive Analysis and text analysis in Systematic Literature Review: A Review of Master Data Management” written by F. Haneem et al. in 2017 [6], “Using Online Reviews for Customer Sentiment Analysis”[8] written by Rae Yule Kim in 2021, and “Sentiment Analysis on Product Review Data” [10] by Xing Fang. et al.

Data Preparation and Cleaning

For data preparation and cleaning of the text data was done using references from “Sentiment Analysis using Support Vector Machines”[5] by A. Gupta et al., “A Comparative Evaluation of Traditional Machine Learning and Deep Learning Classification Techniques for Sentiment Analysis” [11] written by “K.Dhola et al., “Support Vector Machine and Convolutional Neural Network Approach to Customer Review Sentiment Analysis”[15] by Uma et al., “Application of Machine learning techniques to sentiment analysis” [14] by Jain et al. These papers highlight the pre-processing steps of preparing the text data and feature extraction to maximize the accuracy of classifiers, which include but are not limited to the tokenization, stop word removal, and stemming process.

Articles on the models for Sentiment Analysis

After data exploration, the next step was to decide on which algorithms I would choose to conduct sentiment analysis with. According to “A Review On Sentiment Analysis Methodologies, Practices, And Applications”[13] by Mehta et al. which conducted a comparative study of techniques used within sentiment analysis articles, machine learning methods, such as SVM, Naive Bayes, and neural networks have the highest accuracy and can be considered as the baseline learning methods. I will also be using a pre-trained BERT model for my deep learning model. As training a BERT model may take up to several weeks to train, I will be using a PyTorch pre-trained BERT model, similar to "Aspect-Based Sentiment Analysis using Bert" [7] by Hoang et al. which also utilized a pre-trained model for their project.

N-grams

Numerous sentiment analysis studies have been done using different N-grams (Unigrams, Bigrams, Trigrams) on the same dataset. In "Sentiment Analysis using Support Vector Machine"[3] by Zainuddin et al. the paper compared the effectiveness of Unigrams, Bigrams, and Trigrams in training an SVM sentiment classifier. They found that for an SVM, Unigrams greatly outperformed bigrams and trigrams. A similar conclusion was found in "Performance Study of N-grams in the Analysis of Sentiments"[9] by Ojo et al. which compared the accuracy of different N-grams with traditional machine learning techniques and deep learning techniques, they found that unigrams seemed to provide the best accuracy across all algorithms. I expect that Unigrams would also provide the best accuracy in this project.

Descriptive Analytics | Exploratory Data Analysis

My dataset consists of JSON files that contained information on the reviews. The dataset contains a total of 6,990,280 reviews, of which there is a star rating, the review text, the date, the ID of the business being reviewed, and review tags (Useful, Funny, Cool). The dataset is not balanced as there are almost as many 5-star reviews as all the other stars combined. Below is some of the exploratory data analysis of the Yelp Open dataset.

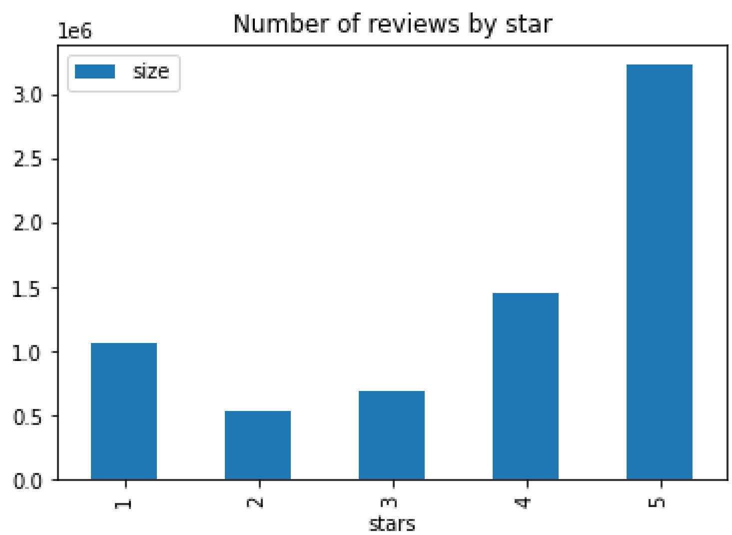
Data Acquisition

The data from this project is a portion of data from Yelp as part of its Yelp Open Dataset. The Yelp Open dataset is open-source and released by Yelp as a part of its Yelp Dataset Challenge.

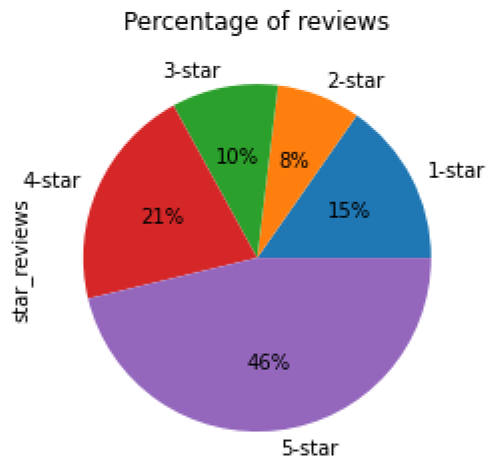
Visualizing the Data

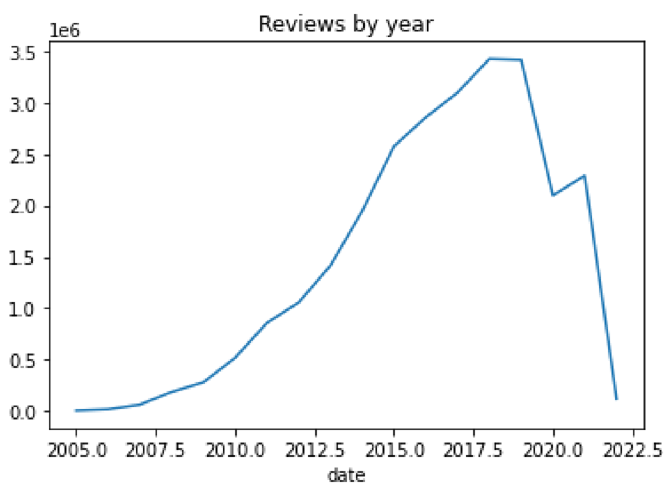
Rating Distribution:

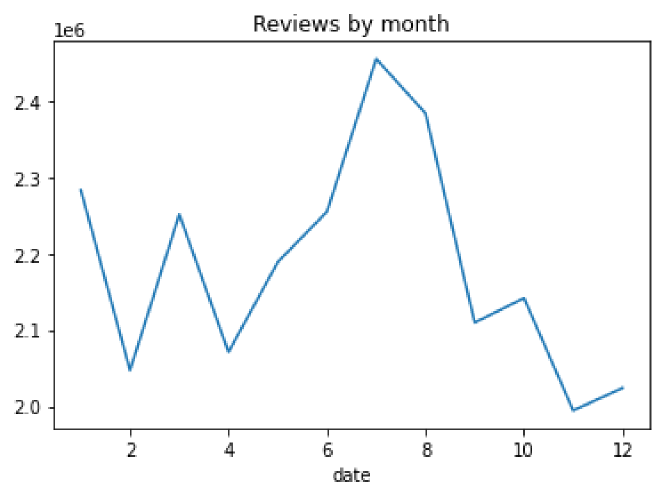
Stars	Reviews	%
1	1’069’561	15%
2	544’240	8%
3	619’934	10%
4	1’452’918	21%
5	3’231’627	46%



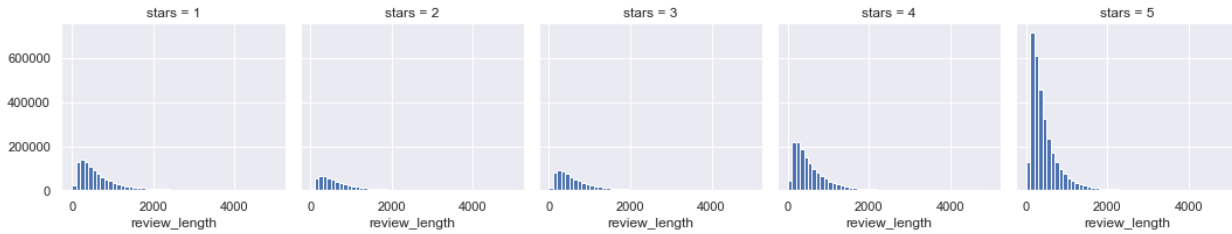
There are much more 5-star reviews within the dataset than other star ratings.







The dataset contains reviews from 2005 to 2022, with most reviews being submitted in the summer months of June to August, although the difference is not large.



The above graph depicts the review length of the reviews at the respective star rating. The reviews seem to follow the trend where most reviews are between 0 and 500, the 5 stars graph look much different from the other stars due to the number of reviews that are 5-stars rating.

Correlation table between review length and review tags (Useful, Funny, Cool)

Correlation table	Useful	Funny	Cool	Review length
Useful	1	0.768755	-0.729256	0.799799
Funny	0.768755	1	-0.485747	0.960635
Cool	-0.729256	-0.485747	1	-0.695158
Review length	0.799799	0.960635	-0.695158	1

There's a negative correlation between *Cool + useful, Cool + Funny, Cool + Length* which implies that cool reviews tend to be short, not useful to others, and unfunny.

And a positive correlation between *Funny + useful, Funny + length, Useful + length*. While longer reviews tend to be funny and useful

Methodology and Experiments

Aim of the study

The aim of this study is to train and compare machine learning algorithms (Naïve Bayes, Logistic Regression, Random Forest, Support Vector Machines) with a deep learning algorithm (BERT) on two tasks. The first is to evaluate a given Yelp text review on its sentiment, positive or negative, and secondly, to predict the text's star rating based on the review text. The machine learning models used in this paper will be trained on review texts converted using TF-IDF vectorizer. I have opted to utilize a pre-trained BERT model, as the process of training such a model from scratch would demand significant computational resources and extensive time.

Selection of Response Variable

In our experiment, the response variable would be “0” (negative) and “1” (positive) sentiment for the sentiment classifier, and a star rating of 1 to 5 for the rating prediction classifier. The following were used as inputs to train the models:

- Text of the reviews
- Star ratings of the reviews
- Tags of the reviews (Cool, Funny, Helpful)

Choice of Factors and Levels

In this experiment, the factors will be the different algorithms and the selection of hyperparameters. The factors could also be different sampling approaches as the dataset is highly imbalanced. The sampling approach I will be using is to resample the same number of training examples from each star rating, in order to avoid bias.

Choice of Experiment Design

Text Processing

The Yelp open dataset contains information about the reviews, not limited to the text of the review. For Machine Learning techniques, I will pre-process the text by removing stop words, numbers, and punctuation. Followed by lemmatization, and TF-IDF calculations. For the Deep Learning model BERT I will be giving the full text to preserve the original sentence structure.

TF-IDF Vectorizer

TF-IDF is a product of term-frequency tf(t, d) and inverse document-frequency idf(t):

tf-idf(*t*, *d*) = tf(*t, d*) × idf(*t*)

Where tf(t,d) is the frequency of the term *t* in document *d* and idf(t) is defined as:

$$idf(t) = \log \frac{n}{1 + df(t)}$$

Where n is the total number of documents in this document set and df(t) is the number of documents in the document set that contains term t.

Randomization (Train/Test Split)

I will split the data into training, and test sets into a 70:30 split to train and evaluate the models.

Dataset Splitting

As the raw data is imbalanced, I will be reconstructing a balanced training dataset by resampling the raw data for a training dataset with 500’000 reviews from each star rating (1-5 stars). It will be mainly under-sampling 5-star reviews as they massively outnumber the other reviews.

Experiment

With a balanced training set by undersampling the 5-star reviews, I want to evaluate the performance of the algorithms in predicting the positive/negative sentiment of a text review compared to predicting the specific star rating of a given text review. The classifiers will be Naive Bayes, Logistic

Regression, Random Forest, Support Vector Machines, and Google's pre-trained BERT model made available by PyTorch. I will be measuring the performance of the models through their accuracy, F1 score, and confusion matrix.

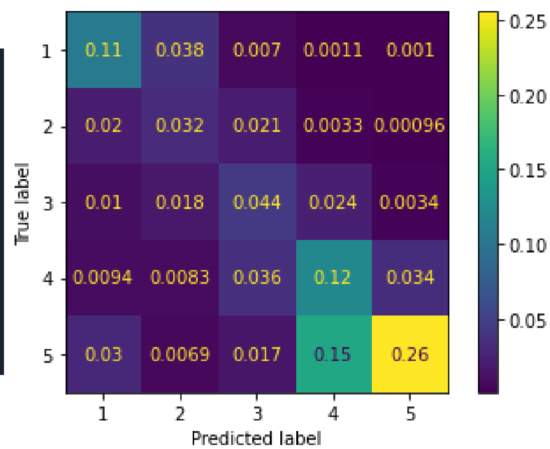
Results

The following presents a classification report and the confusion matrix of the various machine learning algorithms applied to a test dataset. A comparison is made between between simply predicting a positive/negative sentiment, and predicting the specific star rating of a review.

The models were trained by undersampling 400'000 samples from each star rating so as to avoid training bias. However, the testing was done on randomly selected 500'000 samples from the entirety of the data. The only exception was predicting 5-star classification on the BERT model, due to its high computational demands, the sample size was reduced in order to prevent the Google Colab machine from crashing.

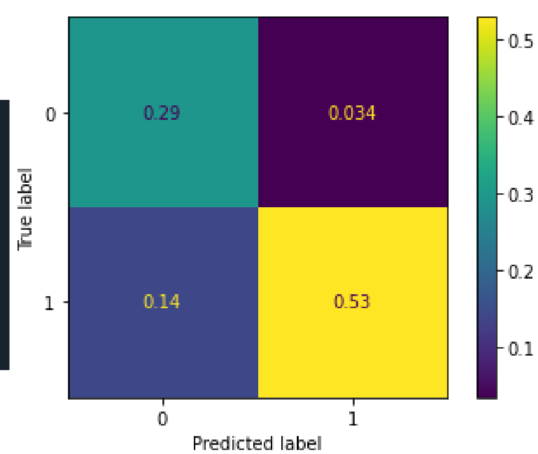
<u>5-Star</u> classification Naive Bayes:

```
Testing Set
Accuracy:  0.556568

              precision    recall  f1-score   support

           1     0.6027    0.6902    0.6435     76303
           2     0.3072    0.4074    0.3503     38783
           3     0.3512    0.4451    0.3926     49443
           4     0.3982    0.5767    0.4711    104155
           5     0.8651    0.5523    0.6741    231316

    accuracy                         0.5566    500000
   macro avg     0.5049    0.5343    0.5063    500000
weighted avg     0.6337    0.5566    0.5742    500000
```



<u>Sentiment classification</u> Naive Bayes:

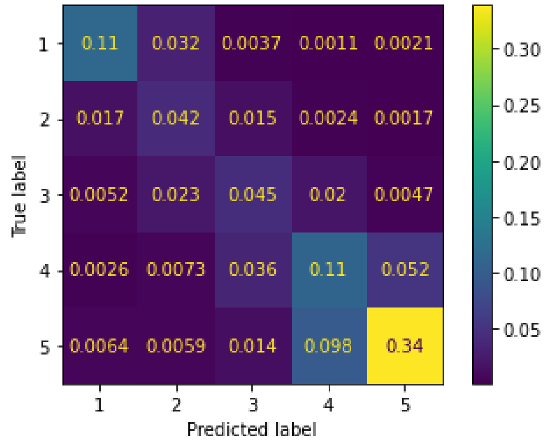
```
Testing Set
Accuracy:  0.82453

              precision    recall  f1-score   support

           0     0.6760    0.8965    0.7708    164529
           1     0.9396    0.7893    0.8579    335471

    accuracy                         0.8245    500000
   macro avg     0.8078    0.8429    0.8143    500000
weighted avg     0.8528    0.8245    0.8292    500000
```



<u>5-Star</u> Logistic Regression:

Testing Set
Accuracy: 0.650252

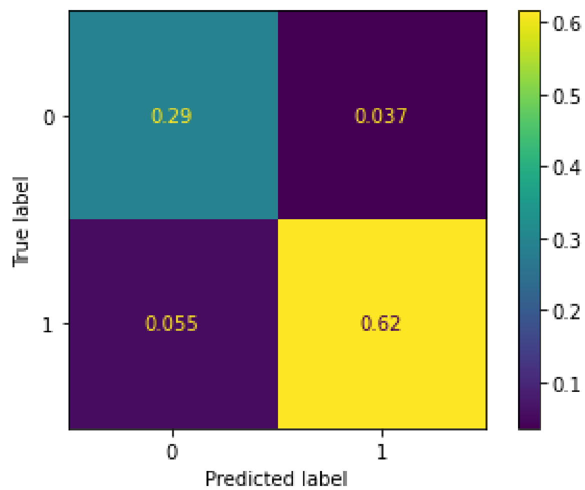
	precision	recall	f1-score	support
1	0.7860	0.7471	0.7661	76303
2	0.3803	0.5403	0.4464	38783
3	0.3998	0.4593	0.4275	49443
4	0.4753	0.5279	0.5002	104155
5	0.8479	0.7326	0.7860	231316
accuracy			0.6503	500000
macro avg	0.5778	0.6014	0.5852	500000
weighted avg	0.6802	0.6503	0.6617	500000



<u>Sentiment classification</u> Logistic Regression:

Testing Set
Accuracy: 0.908618

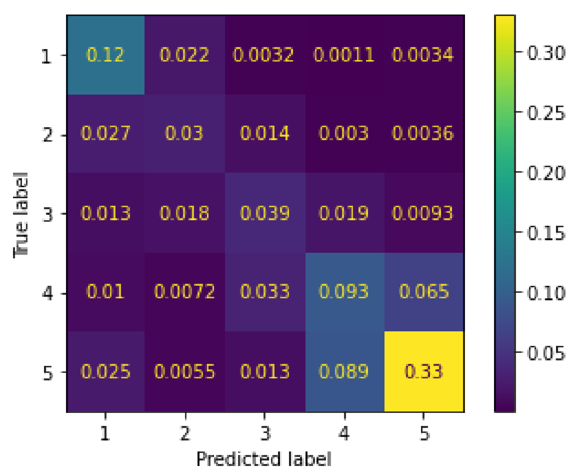
	precision	recall	f1-score	support
0	0.8422	0.8888	0.8649	164529
1	0.9439	0.9183	0.9310	335471
accuracy			0.9086	500000
macro avg	0.8931	0.9036	0.8979	500000
weighted avg	0.9105	0.9086	0.9092	500000



<u>5-Star</u> Random Forest:

Testing Set
Accuracy: 0.615466

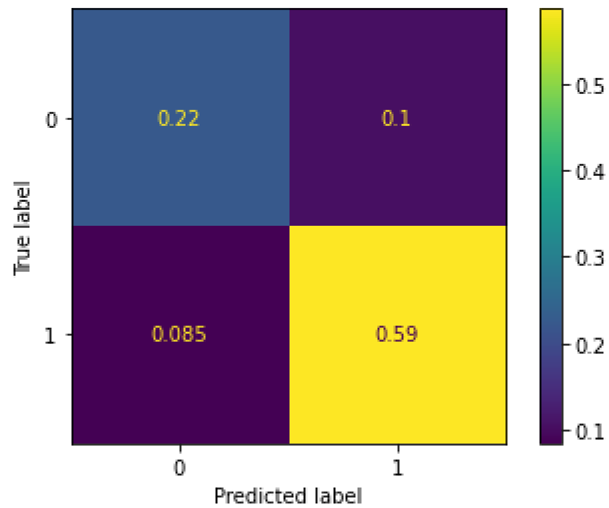
	precision	recall	f1-score	support
1	0.6197	0.8061	0.7007	76303
2	0.3630	0.3822	0.3724	38783
3	0.3815	0.3989	0.3900	49443
4	0.4536	0.4472	0.4504	104155
5	0.8032	0.7138	0.7559	231316
accuracy			0.6155	500000
macro avg	0.5242	0.5496	0.5339	500000
weighted avg	0.6266	0.6155	0.6179	500000



<u>Sentiment classification</u> Random Forest:

Testing Set
Accuracy: 0.810738

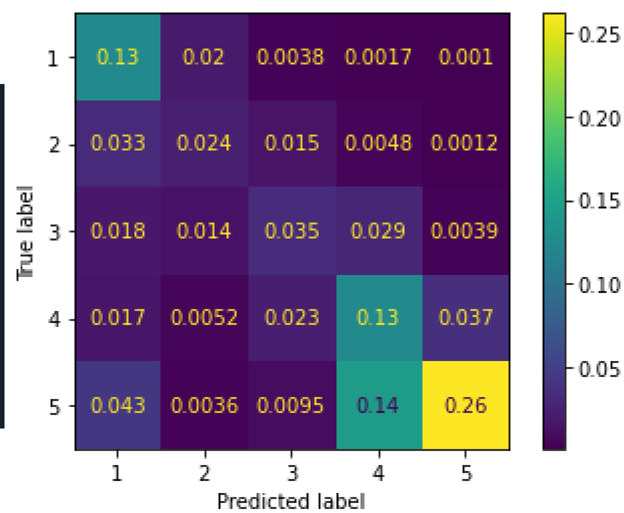
	precision	recall	f1-score	support
0	0.7263	0.6818	0.7033	164529
1	0.8485	0.8740	0.8610	335471
accuracy			0.8107	500000
macro avg	0.7874	0.7779	0.7822	500000
weighted avg	0.8083	0.8107	0.8091	500000



<u>5-star</u> SVM

Testing Set
Accuracy: 0.572264

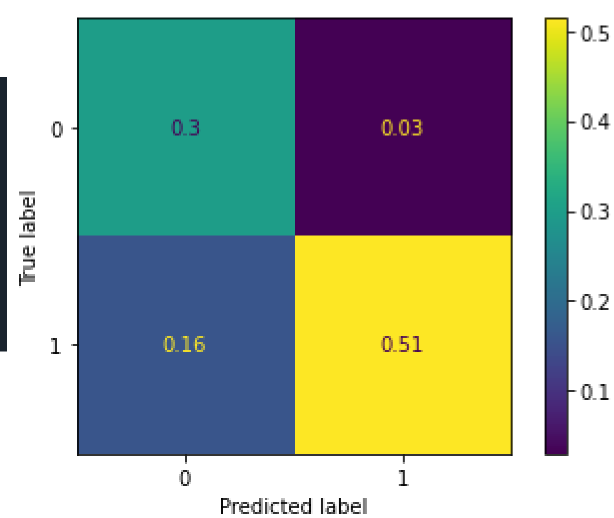
	precision	recall	f1-score	support
1	0.5307	0.8267	0.6465	76303
2	0.3596	0.3055	0.3304	38783
3	0.3998	0.3497	0.3731	49443
4	0.4128	0.6052	0.4908	104155
5	0.8596	0.5658	0.6825	231316
accuracy			0.5723	500000
macro avg	0.5125	0.5306	0.5046	500000
weighted avg	0.6321	0.5723	0.5791	500000



<u>Sentiment classification</u> SVM:

Testing Set
Accuracy: 0.813022

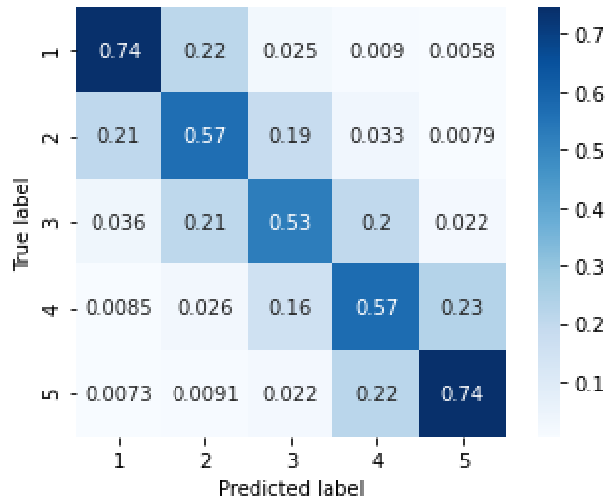
	precision	recall	f1-score	support
0	0.6559	0.9084	0.7618	164529
1	0.9446	0.7662	0.8461	335471
accuracy			0.8130	500000
macro avg	0.8003	0.8373	0.8039	500000
weighted avg	0.8496	0.8130	0.8184	500000



5-Star BERT:

Testing Set
Accuracy: 0.658952

	precision	recall	f1-score	support
0	0.7675	0.7444	0.7557	31175
1	0.4480	0.5670	0.5005	22635
2	0.4892	0.5274	0.5076	30996
3	0.5437	0.5731	0.5580	62024
4	0.8319	0.7444	0.7857	103170
accuracy			0.6590	250000
macro avg	0.6161	0.6313	0.6215	250000
weighted avg	0.6751	0.6590	0.6652	250000

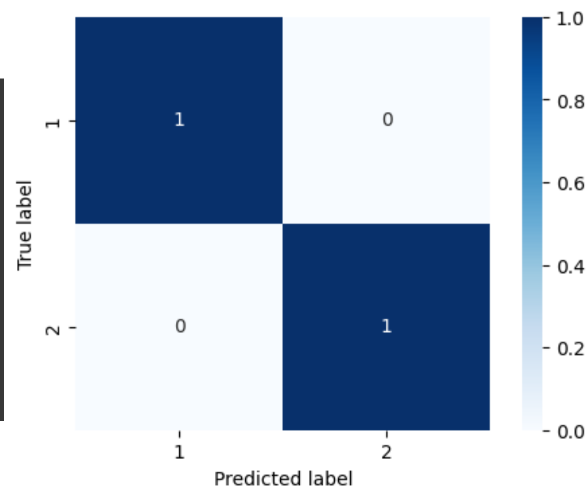


The 5-star BERT was evaluated on a dataset of 250'000 reviews instead of 500'000 due to insufficient computational resources.

Sentiment classification BERT:

testing Set
Accuracy: 1.0

	precision	recall	f1-score	support
0	1.0000	1.0000	1.0000	164529
1	1.0000	1.0000	1.0000	335471
accuracy			1.0000	500000
macro avg	1.0000	1.0000	1.0000	500000
weighted avg	1.0000	1.0000	1.0000	500000



Most surprisingly the BERT model was able to predict sentiment with perfect accuracy, I had expected at least a few reviews to be incorrectly classified, but no errors were made.

All algorithms consistently perform the best when accessing 5-star and 1-star reviews, with less accuracy observed for 2-4 star reviews. This may be due to the more distinctive language used in 1-star and 5-star reviews, compared to the language used in 2-4 star reviews, which is more similar.

While all machine learning algorithms are quite accurate in predicting sentiment, the BERT model excels in predicting sentiment with perfect accuracy. Furthermore, in predicting the 5-star rating, it demonstrates significantly higher accuracy for 2-4 star reviews, a task with which other machine learning algorithms struggle.

Accuracy of algorithms on the test dataset for 5-star rating accuracy

Naive Bayes	Logistic Reg.	Random Forests	SVMs	BERT
0.556568	0.650252	0.615466	0.572264	0.658952

Accuracy of algorithms on the test dataset for Sentiment detection accuracy

Naive Bayes	Logistic Reg.	Random Forests	SVMs	BERT
0.82453	0.908618	0.810738	0.813022	1.0

Discussion

The outcomes of this paper corroborate the sentiment that deep learning models, especially transformer architectures like BERT, offer a significant edge in handling complex natural language processing tasks. The architecture's ability to understand context and semantic nuances appears to be central to its success in our experiments. However, the 100% accuracy in the BERT's sentiment classification does invite some scrutiny. It’s worth exploring in future research whether such a performance can be replicated across other datasets or if there’s something unique about the Yelp Open dataset that led to this outcome.

While undersampling the predominant 5-star reviews in the training dataset made the dataset less biased, it may have excluded some more nuanced insights that could’ve been found within the unused reviews. The challenges faced by the machine learning models in predicting 2-4 star reviews might indicate that the reviews have some subtleties that the models find challenging in classifying. It might be because consumers blend positive and negative sentiments together within the same review, making them harder to classify. While the BERT’s accuracy is much better, the resources and time required are significantly higher compared to the machine learning models. While not considered in this project, the computational cost, resulting in a trade-off between accuracy and computational efficiency might become crucial when considering real-world applications such as businesses or social media platforms that process large volumes of reviews in real-time.

Conclusions and Future Work

In this paper, we sought to utilize machine learning and deep learning techniques to conduct sentiment analysis on Yelp business reviews. This project examines the effectiveness of Naïve Bayes, Logistic Regression, Random Forest, Support Vector Machines, and BERT models on the Yelp open dataset which includes 6,990,280 reviews. Each model’s performance is assessed based on its classification report and confusion matrices.

The objectives were to determine the sentiment (positive or negative) of a given review and to predict the specific star rating of the review text. Traditional machine learning models (Naive Bayes, Logistic Regression, Random Forests, and Support Vector Machines) showcased notable accuracy rates, particularly with more clear-cut sentiment reviews such as 1-star and 5-star reviews. However, their performance was worse and varied in predicting more nuanced sentiments as seen in their results in classifying 2-4 star reviews. The deep learning model used for this paper, BERT showed incredible proficiency in sentiment classification achieving perfect accuracy in sentiment classification, and noticeably superior performance in predicting specific star ratings for 2-4 star reviews in comparison to traditional machine learning models.

For future works, it would be a good idea to repeat the process for another type of data, as Yelp is composed of crowd-sourced reviews about businesses, the results found in this paper might not generalize well into other platforms such as travel blogs, e-commerce platforms, or movie reviews due to different terminology and distinct sentiment nuances in each dataset. While this paper used undersampling to address the dataset imbalances, exploring alternative strategies such as oversampling or synthetic data augmentation might help with the generalizability of the findings in this project. This project could also be repeated with newer deep learning architectures or BERT derivatives, such as RoBERTa, XLNet, or DistilBERT.

Appendix - A | Data Source & GitHub Link

Github Link

https://github.com/Tusk98/MRP

Data Source

https://www.yelp.com/dataset

References

[1]Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. (2002). “Thumbs up?: sentiment classification using machine learning techniques”. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing, July 28-31, 2002, Philadelphia, Pennsylvania, USA, 79-86

[2] Andrew Elkouri (2015), “Predicting the Sentiment Polarity and Rating of Yelp Reviews” arXiv:1512.06303

[3] N. Zainuddin and A. Selamat, "Sentiment analysis using Support Vector Machine," 2014 International Conference on Computer, Communications, and Control Technology (I4CT), Langkawi, Malaysia, 2014, pp. 333-337, doi: 10.1109/I4CT.2014.6914200. https://ieeexplore.ieee.org/abstract/document/6914200

[4] Alec Go, Richa Bhayani, and Lei Huang. (2009). “Twitter Sentiment Classification using Distant Supervision”.https://www.researchgate.net/publication/228523135_Twitter_sentiment_classification_using_distant_supervision

[5] A. Gupta, P. Tyagi, T. Choudhury and M. Shamoon, "Sentiment Analysis Using Support Vector Machine," 2019 International Conference on contemporary Computing and Informatics (IC3I), Singapore, 2019, pp. 49-53, doi: 10.1109/IC3I46837.2019.9055645.

[6] F. Haneem, R. Ali, N. Kama and S. Basri, "Descriptive analysis and text analysis in Systematic Literature Review: A review of Master Data Management," (2017) International Conference on Research and Innovation in Information Systems (ICRIIS), Langkawi, Malaysia, 2017, pp. 1-6, doi: 10.1109/ICRIIS.2017.8002473.

[7] Mickel Hoang, Oskar Alija Bihorac, and Jacobo Rouces. (2019). “Aspect-Based Sentiment Analysis using BERT”. In Proceedings of the 22nd Nordic Conference on Computational Linguistics, pages 187–196, Turku, Finland. Linköping University Electronic Press.

[8] Rae Yule Kim (2021), "Using Online Reviews for Customer Sentiment Analysis", in IEEE Engineering Management Review, vol. 49, no. 4, pp. 162-168, 1 Fourth quarter, Dec. 2021, doi:10.1109/EMR.2021.3103835.

[9] Ojo, Olumide & Gelbukh, Alexander & Calvo, Hiram & Adebanji, Olaronke. (2021). Performance Study of N-grams in the Analysis of Sentiments. Journal of the Nigerian Society of Physical Sciences. 3. 477-483. 10.46481/jnsps.2021.201.

[10] Fang, Xing. Zhang, Justin. “Sentiment Analysis using product review data” Journal of Big Data 2, 5 (2015). https://doi.org/10.1186/s40537-015-0015-2

[11] K. Dhola and M. Saradva, "A Comparative Evaluation of Traditional Machine Learning and Deep Learning Classification Techniques for Sentiment Analysis," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 932-936, doi: 10.1109/Confluence51648.2021.9377070.

[12] S. Singh, K. Kumar and B. Kumar, "Sentiment Analysis of Twitter Data Using TF-IDF and Machine Learning Techniques," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 252-255, doi: 10.1109/COM-IT-CON54601.2022.9850477.

[13] P. Mehta, S. Pandya, “A Review On Sentiment Analysis Methodologies, Practices And Applications” ISSN 2277-8616

[14] A. P. Jain and P. Dandannavar, "Application of machine learning techniques to sentiment analysis," 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Bangalore, India, 2016, pp. 628-632, doi: 10.1109/ICATCCT.2016.7912076.

[15] R. Uma, A. S. H, P. Jawahar and B. V. Rishitha, "Support Vector Machine and Convolutional Neural Network Approach to Customer Review Sentiment Analysis," 2022 1st International Conference on Computational Science and Technology (ICCST), CHENNAI, India, 2022, pp. 239-243, doi: 10.1109/ICCST55948.2022.10040381.