UTM
UNIVERSITI TEKNOLOGI MALAYSIA

UNIVERSITI TEKNOLOGI MALAYSIA

Research Design and Analysis in Data Science
MCST1043

Sentiment Analysis of Amazon Reviews Using Machine Learning

Done by: Omar Mohammed Ali Albaagari (MCS241021)
Prepared for: Prof. Madya.Ts.Dr. Mohd Shahizan bin Othman

Innovating Solutions

OUTLINES

1 Introduction

2 Literature Review

3 Methodology

4 Initial Results

5 Discussion and Future Work

6 References

INTRODUCTION

- Emotions are present in every single situation in which people engage with one another (De Saa and Ranathunga, 2020).

- In the twenty-first century, the internet has grown into a technology that is indispensable to our everyday life (Ripa et al., 2021).

- People are purchasing things from a large number of e-commerce websites in the current period, and it is more likely that they would first assess the products before purchasing them (Rathor et al., 2018).

- Sentimental analysis is one of the machine learning processing techniques that helps detect feelings (Rajat et al., 2021).

INTRODUCTION

- This approach enables business owners to collect information about the perspectives of their customers via various online media, such as social media, and analyses of websites that allow for online shopping.
- Sentiment analysis represents the behaviour of the consumer with regard to the product, as well as the reputation of the company.

Problem Statement

- Customer ratings and reviews reflect buyer judgment but may not always convey true sentiment.
- Businesses face challenges in accurately assessing customer satisfaction based solely on star ratings.
- There is a need for advanced sentiment analysis techniques to interpret hidden emotions in customer reviews.

Research Goal

- To identify patterns in Office Products in Amazon to enhance the understanding of customers behavior by utilizing VADER and Roberta techniques in Office Products reviews

Research Questions	Research Objectives
a) What preprocessing steps to carry out for the analyzing sentiment analysis from office product dataset?	a) To conduct a preprocessing of the office products reviews datasets for sentiment analysis.
b) What relevant keywords can be identified and retrieved by VADER and Roberta from the review's dataset?	b) To train a machine learning model that is capable of sorting customer evaluations into three unique sentiment categories, namely positive, neutral, and negative categories?
c) What conclusions may be derived from the customers purchases?	c) To develop a dashboard that summarize the analysis and making conclusion of their behavior.

Literature Reviews

Several distinct degrees of investigation have been conducted on the subject of sentiment analysis. It is largely possible to identify sentiments and viewpoints at the level of the text, phrase, or aspect (Do et al., 2019). An illustration of the degrees of sentiment analysis may be seen in Figure 2.1. The first two levels are very great to go through, but they are also really challenging. In spite of this, the third level is more challenging than the levels that came before it since it demands a more comprehensive examination. (Cambria et al., 2017).

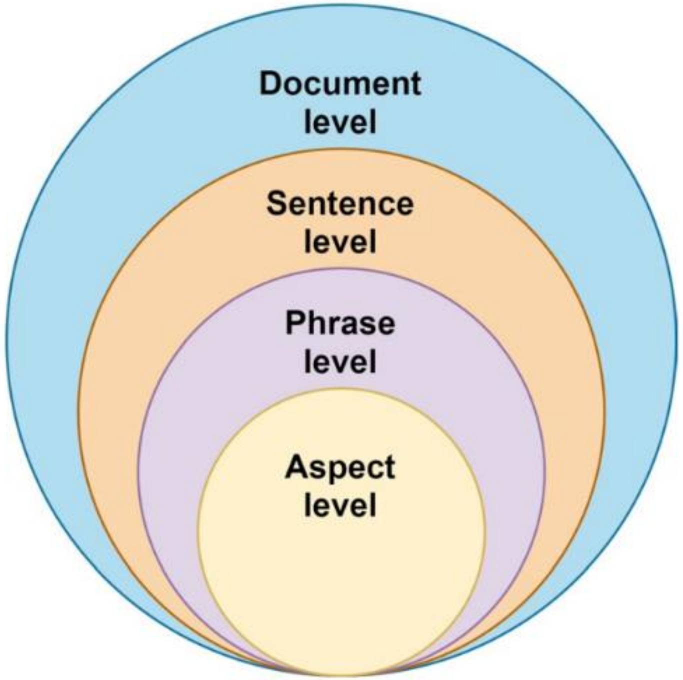

Figure 2.1 Various levels of emotional analysis (Wankhade et al., 2022).

Literature Review

Title	Author	Finding
A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends	Birjali, M., Kasri, M., and Beni-Hssane, A. (2021)	Reviewed various sentiment analysis methods, highlighting traditional and deep learning models. Discussed challenges in sarcasm detection and trends in multilingual data and real-time sentiment tracking.
Sentiment Analysis Based on Deep Learning: A Comparative Study	Dang, N. C., Moreno-García, M. N., and De la Prieta, F. (2020)	Found that deep learning models like LSTM and GRU outperform traditional machine learning models, especially with word embeddings.
Returning the N to NLP: Towards Contextually Personalized Classification Models	Flek, L. (2020)	Emphasized the importance of contextually personalized NLP models by integrating user-specific data, improving tasks like sentiment analysis.
Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models	Gope, J. C., Tabassum, T., Mabrur, M. M., Yu, K., and Arifuzzaman, M. (2022)	Demonstrated that LSTM and CNN models outperform traditional machine learning models for sentiment analysis, with word embeddings enhancing performance.

Literature Review

Title	Author	Finding
Sentiment Analysis via Semi-Supervised Learning: A Model Based on Dynamic Threshold and Multi-Classifiers	Han, Y., Liu, Y., and Jin, Z. (2020)	Proposed a semi-supervised model using dynamic thresholding and multiple classifiers, improving accuracy with limited labeled data.

Research Methodology

Data Collection

- A dataset consisting of reviews of office products was gathered from the Amazon Reviews Repository collection [Amazon Reviews'23]. There are a total of 200,000 rows

	parent_asin	user_id	helpful_vote	asin	text	timestamp	images	verified_purchase	title	rating
0	B01MZ3SD2X	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B01AHHL4X2	Lovely ink. Writes well. The right amount of w...	1677939345945	[]	True	Pretty & I love it!	5.0
1	B08L6H23JZ	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B08L6H23JZ	Overall I'm pretty happy with this purchase bc...	1677939160682	[]	True	2 excellent 1 extremely dry (blue)	4.0
2	B07JDZ5J46	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	2	B07JDZ5J46	[[VIDEOID:63276c19932aa4f3687042b8b9f8613c]] U...	1660188831933	[]	True	I don't get the reviews. Mine are garbage.	1.0
3	B07BR2PBJN	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B004MNX7EW	It's a beautiful color, but even though it had...	1659806066713	[{'small_image_url': 'https://m.media-amazon.c...	True	Ordering Ink online: never a good idea I guess.	4.0
4	B097SFY5ZS	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B019YLRFFS	Idk if I just got a bad batch which is possibl...	1659799390978	[]	True	Mine are iffy at best.	3.0

Figure 3.2 Review Dataset

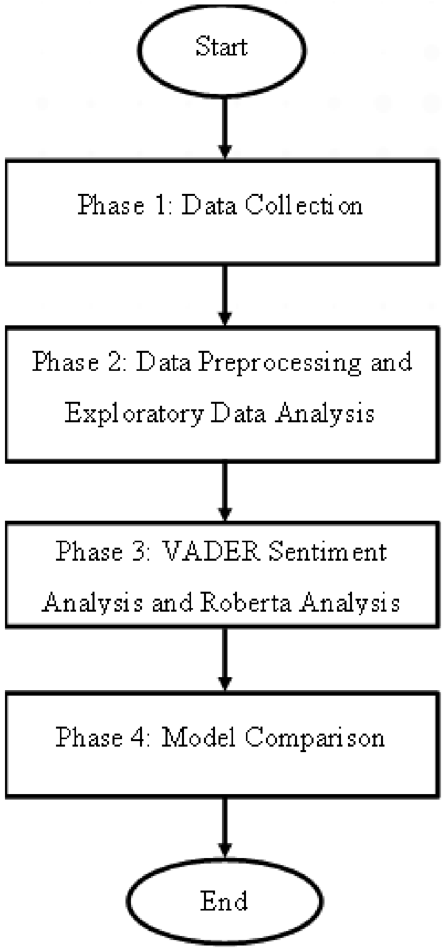

Figure 3.1 Overall research methodology

Research Methodology

Data Preprocessing and Exploratory Data Analysis

Data Cleaning

- Remove any rows with missing product review data or ratings.
- Replace missing ratings (NaN) with the average of other ratings or remove the rows, as there's a large dataset.

Preprocess text by:

- Converting to lowercase
- Removing extra whitespaces
- Replacing digits and punctuation with spaces
- Eliminating extra spaces and tabs
- Tokenize the text into words or sub-words and apply stemming for word standardization.

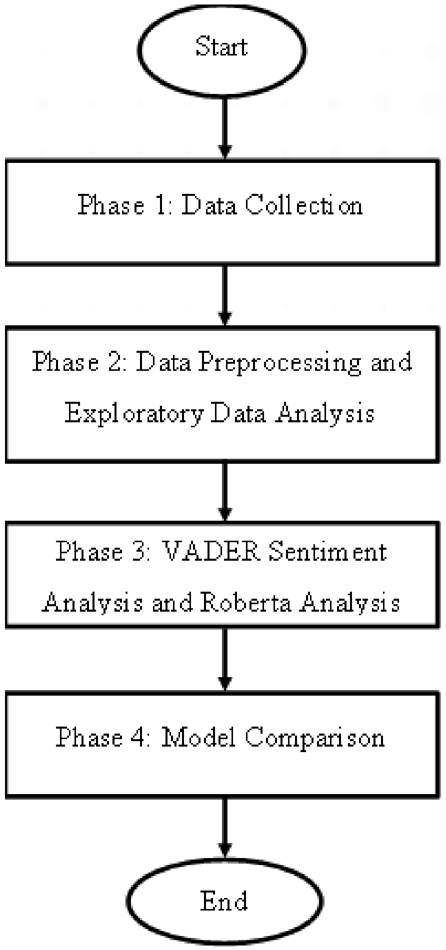

Figure 3.1 Overall research methodology

Research Methodology

<u>VADER and Roberta Sentiment Analysis</u>

VADER for quick, rule-based sentiment analysis, providing polarity scores from -1 (negative) to +1 (positive), ideal for short texts. To capture deeper context, fine-tuned the RoBERTa model, a transformer-based framework with a stronger grasp of word relationships and subtle sentiment shifts. Combining VADER's speed with RoBERTa's contextual accuracy enhances overall sentiment classification.

<u>Model Comparison</u>

After implementing NLP, evaluate the model using accuracy, precision, recall, and F1-score, which balances precision and recall to assess sentiment classification. Finally, select the most accurate model.

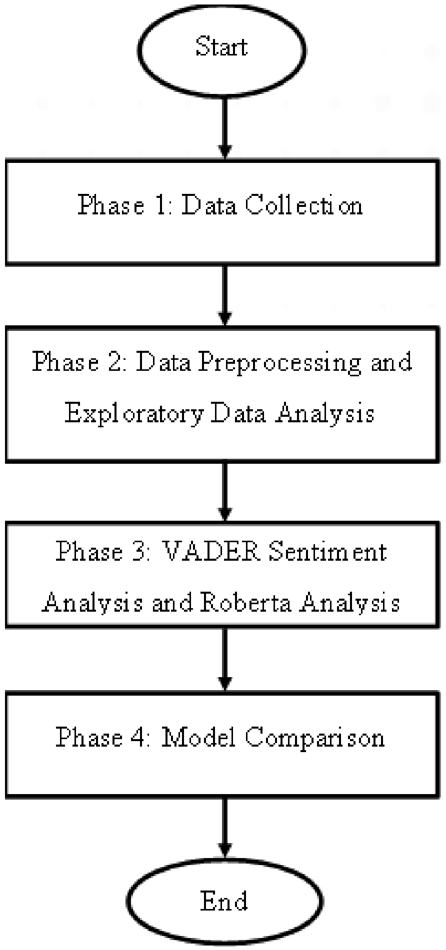

Figure 3.1 Overall research methodology

Exploratory Data Analysis

In figure 4.3 shows the dataset information of each column also the type of the data that used. It can be seen that all columns are non-null, consisting of 6 objects, 3 int64, 1 bool, and 1 float64.

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200000 entries, 0 to 199999
Data columns (total 10 columns):
 #   Column             Non-Null Count   Dtype
---  ------             --------------   -----
 0   parent_asin        200000 non-null  object
 1   user_id            200000 non-null  object
 2   helpful_vote       200000 non-null  int64
 3   asin               200000 non-null  object
 4   text               199975 non-null  object
 5   timestamp          200000 non-null  int64
 6   images             200000 non-null  object
 7   verified_purchase  200000 non-null  bool
 8   title              199964 non-null  object
 9   rating             200000 non-null  float64
dtypes: bool(1), float64(1), int64(2), object(6)
memory usage: 13.9+ MB
```

```
df.columns
```

```
Index(['parent_asin', 'user_id', 'helpful_vote', 'asin', 'text', 'timestamp',
       'images', 'verified_purchase', 'title', 'rating'],
      dtype='object')
```

Figure 4.3 Data Information

Exploratory Data Analysis

In figure 4.4 shows the dataset description which is about the basic statistical analysis such as mean, standard deviation, minimum and maximum values

```
df.describe()
```

	helpful_vote	timestamp	rating
count	200000.000000	2.000000e+05	200000.000000
mean	1.108975	1.545500e+12	4.412790
std	10.618726	8.653161e+10	1.111684
min	0.000000	9.587741e+11	1.000000
25%	0.000000	1.482169e+12	4.000000
50%	0.000000	1.558833e+12	5.000000
75%	0.000000	1.614727e+12	5.000000
max	1561.000000	1.679245e+12	5.000000

Figure 4.4 Dataset Description

Exploratory Data Analysis

The positive sentiment word cloud highlighted frequent terms like "easy," "perfect," "love," "great," "nice," "beautiful," "good," "pen," and "better." Phrases like "good quality" and "works great" reflected product satisfaction, while "price" and "happy" indicated approval of affordability.

Figure 4.5 World Cloud of Positive Sentiment

Exploratory Data Analysis

The negative sentiment word cloud highlighted frequent terms like "small," "little," "bad," "waste," and "disappointed," reflecting dissatisfaction. Words like "time," "hard," and "small" indicated frustration with certain office products.

Figure 4.6 World Cloud of Negative Sentiment

Exploratory Data Analysis

Figure below illustrate the word cloud for neutral sentiment reviews. Word cloud analysis illustrated that "use," "color," "wish," "product," were the most frequent used words in the review. These words can represent that the products don’t meet their meet as their expectation.

Figure 4.7 World Cloud of Neutral Sentiment

Exploratory Data Analysis

Figure 4.8 shows the rating distribution (1–5) for office products on Amazon, with most ratings being five. This suggests high product quality and reasonable pricing, reflecting overall customer satisfaction.

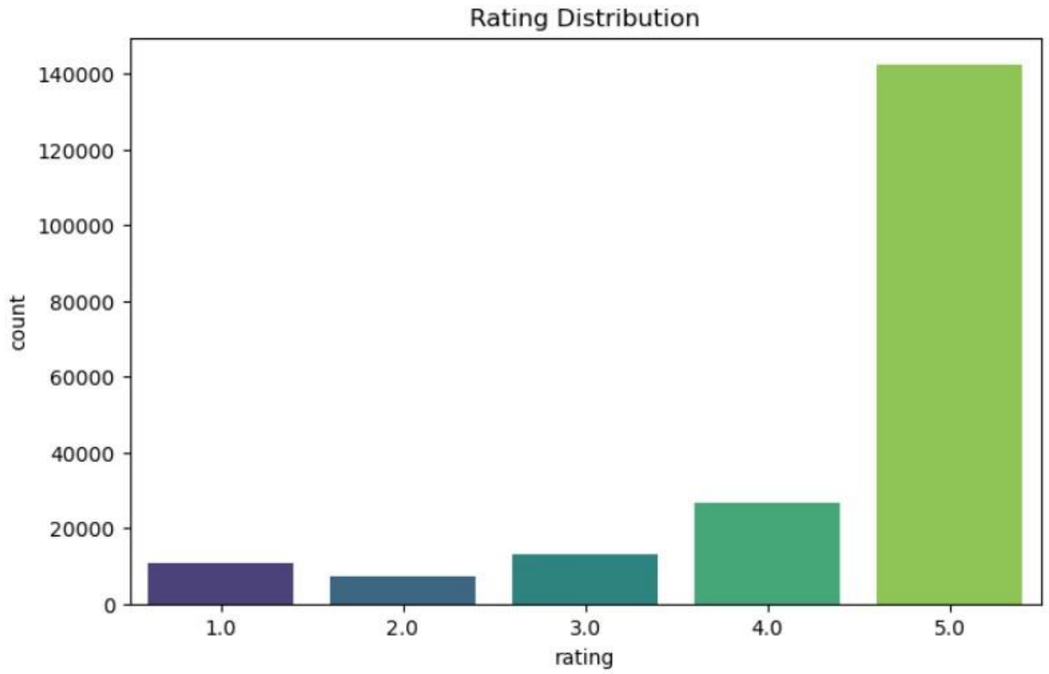

Figure 4.8 Rating Distribution

Exploratory Data Analysis

Figure 4.9 illustrate the verified purchase distribution which indicate the majority of the customers are verified their purchase. Whereas smaller portion are not verifying their purchase

Figure 4.9 Verifies Purchase Distribution

Exploratory Data Analysis

A list of the top ten titles of reviews submitted by consumers is shown in the figure below. "Good Product" accounts for the smallest share, while "Five Star" is the term that receives the most amount of portion

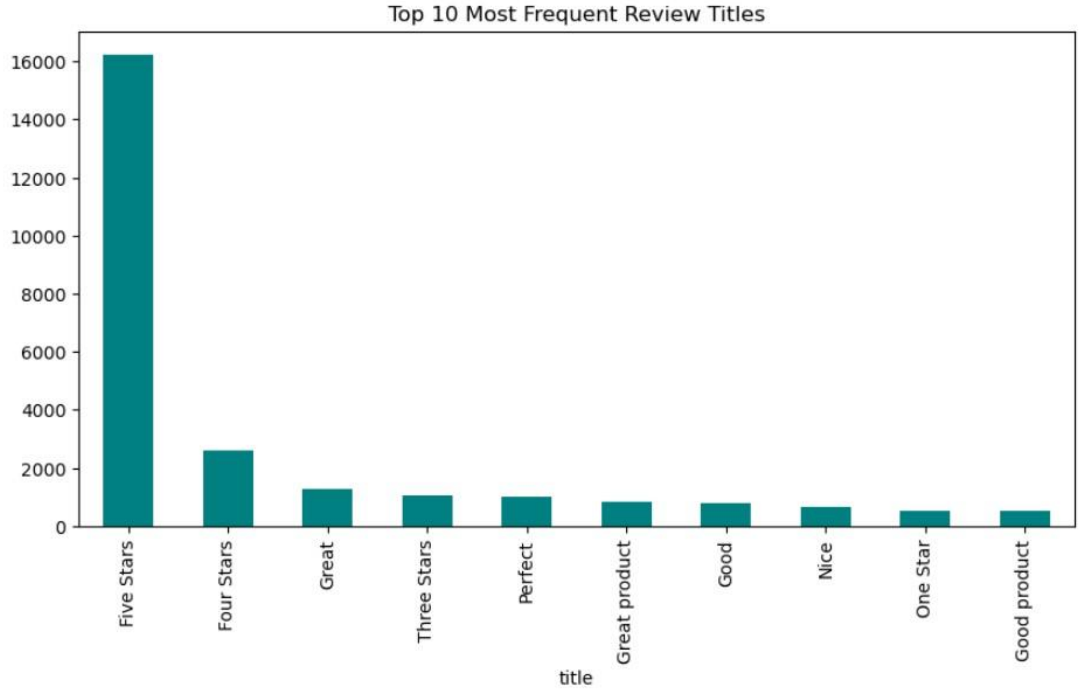

Figure 4.10 Top 10 Most Frequent Review Titles

Exploratory Data Analysis

The scatter plot shows a positive association between ratings and helpful votes, indicating that higher-rated products are more likely to be seen as helpful by customers.

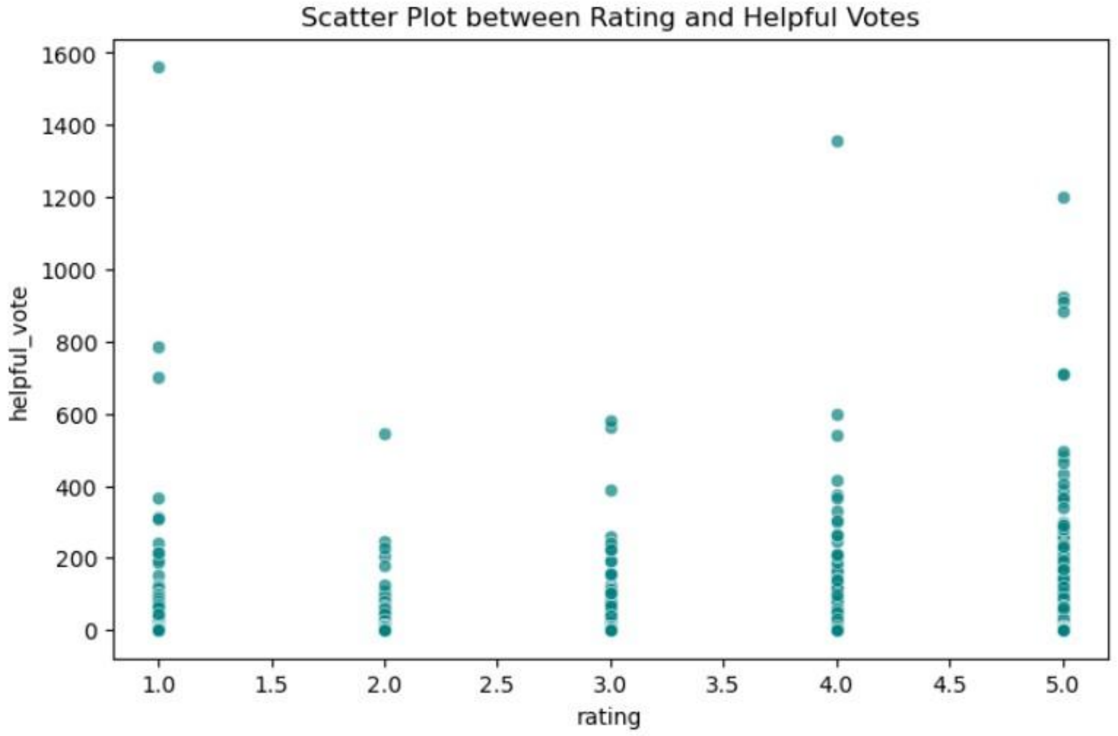

Figure 4.11 Scatter Plot between Rating and Helpful_vote

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Model Development

In this research using two models in the sentiment analysis which are Valence Aware Dictionary and sEntiment Reasoner and Robustly Optimized BERT Pretraining Approach using Python. Some libraries were used including the NLTK, vaderSentiment, Transformers, Pandas, Scikit-learn, and seaborn. Those are helping to approach the sentiment analysis in effective way.

Model Development

Vader Model

	Id	neg	neu	pos	compound	Product_ID	Helpful_Vote	Rating	Time	verified_purchase	Summary	Text
0	1	0.000	0.677	0.323	0.9300	B01MZ3SD2X	0	5.0	1677939345945	True	Pretty & I love it!	Lovely ink. Writes well. The right amount of w...
1	2	0.051	0.771	0.178	0.9481	B08L6H23JZ	0	4.0	1677939160682	True	2 excellent 1 extremely dry (blue)	Overall I'm pretty happy with this purchase bc...
2	3	0.070	0.815	0.115	0.9498	B07JDZ5J46	2	1.0	1660188831933	True	I don't get the reviews. Mine are garbage.	[[VIDEOID:63276c19932aa4f3687042b8b9f8613c]] U...
3	4	0.072	0.755	0.173	0.9941	B07BR2PBJN	0	4.0	1659806066713	True	Ordering Ink online: never a good idea I guess.	It's a beautiful color, but even though it had...
4	5	0.142	0.776	0.082	-0.9306	B097SFY5ZS	0	3.0	1659799390978	True	Mine are iffy at best.	Idk if I just got a bad batch which is possibl...

Figure 4.15 Data Frame of VADER Model

Model Development

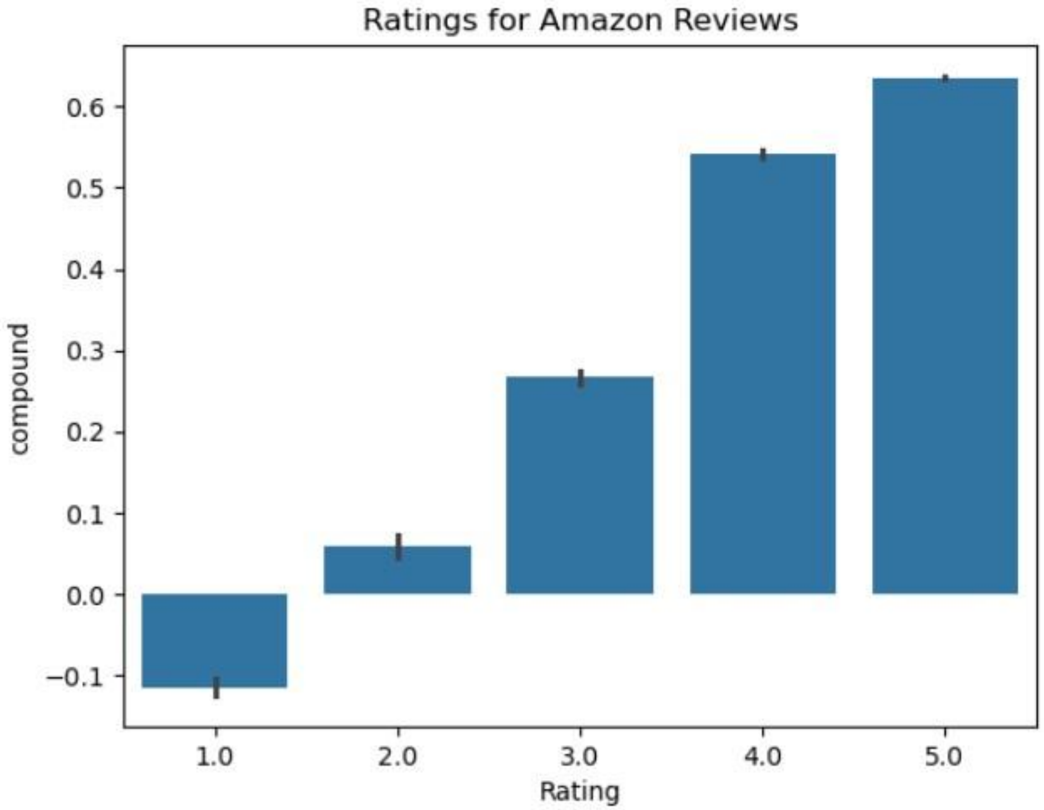

Figure 4.16 Visualization of VADER Sentiment

Model Development

Roberta Model

roberta_neg	roberta_neu	roberta_pos	Product_ID	Helpful_Vote	Rating	Time	verified_purchase	Summary	Text
0.001184	0.016531	0.982284	B01MZ3SD2X	0	5.0	1677939345945	True	Pretty & I love it!	Lovely ink. Writes well. The right amount wet/...
0.066594	0.202709	0.730696	B08L6H23JZ	0	4.0	1677939160682	True	2 excellent 1 extremely dry (blue)	Overall I'm pretty happy purchase bc ink good ...
0.907260	0.081818	0.010923	B07JDZ5J46	2	1.0	1660188831933	True	I don't get reviews. Mine garbage.	[[VIDEOID:63276c19932aa4f3687042b8b9f8613c]] U...
0.156052	0.423542	0.420406	B07BR2PBJN	0	4.0	1659806066713	True	Ordering Ink online: never good idea I guess.	It's beautiful color, even though packed extre...
0.745801	0.219413	0.034786	B097SFY5ZS	0	3.0	1659799390978	True	Mine iffy best.	Idk I got bad batch possible I suppose bc let'...

Figure 4.17 Data frame of Roberta Model

Discussion and Future Work

Achievements for this research are:

- A cleaned dataset can be handled by removing the duplicates and handling the missing values and irrelevant information that occurred in the dataset.
- EDA has been done successfully
- VADER and Roberta models are half way developed

Future Work in this project

(a) Get the sentiment analysis of both VADER and Roberta.

(b) Determine the accuracy of both models to get which one is more accurate of the sentiment analysis

(c) Visualization of the insights from office product reviews through dashboard

References

Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). A practical guide to sentiment analysis (Vol. 5). Springer.

Rathor, A. S., Agarwal, A., & Dimri, P. (2018). Comparative study of machine learning approaches for Amazon reviews. Procedia Computer Science, 132, 1552–1561.

Do, H. H., Prasad, P. W., Maag, A., & Alsadoon, A. (2019). Deep learning for aspect-based sentiment analysis: A comparative review. Expert Systems with Applications, 118, 272–299.

De Saa, E., & Ranathunga, L. (2020). Self-reflective and introspective feature model for hate content detection in Sinhala YouTube videos. In 2020 From Innovation to Impact (FITI) (Vol. 1, pp. 1–6). IEEE.

Flek, L. (2020). Returning the N to NLP: Towards contextually personalized classification models. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 7828–7838).

References

Han, Y., Liu, Y., & Jin, Z. (2020). Sentiment analysis via semi-supervised learning: A model based on dynamic threshold and multi-classifiers. Neural Computing and Applications, 32, 5117–5129.

Khalid, M., Ashraf, I., Mehmood, A., Ullah, S., Ahmad, M., & Choi, G. S. (2020). Gbsvm: Sentiment classification from unstructured reviews using ensemble classifier. Applied Sciences, 10(8), 2788.

Kumar, A., & Garg, G. (2020). Systematic literature review on context-based sentiment analysis in social multimedia. Multimedia Tools and Applications, 79(21), 15349–15380.

Ripa, S. P., Islam, F., & Arifuzzaman, M. (2021). The emerging threat of phishing attacks and the detection techniques using machine learning models. In 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI) (pp. 1–6). IEEE.

References

Gope, J. C., Tabassum, T., Mabrur, M. M., Yu, K., & Arifuzzaman, M. (2022). Sentiment analysis of Amazon product reviews using machine learning and deep learning models. In 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE) (pp. 1–6). IEEE.

THANK YOU

univteknologimalaysia utm.my utmofficial