International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Sentiment Analysis of Amazon Reviews using NLTK Vader and Robert

Jitendra Soni, Ayush Raghuwanshi, Kavleen Kour, Aman Raghuwanshi
Institute of Engineering & Technology , DAVV Indore (M.P.)

Abstract

With the advancement of technology, huge volume of data is exchanged every-day, example - from social media sites like Instagram, Twitter, or e-commerce sites like Amazon, Flipkart etc. The human brain is unable to process this large amount of data (mainly out of which is unstructured) and this results in the development of technology which do the work of humans. One such methodology is Sentiment Analysis.

Sentiment analysis plays a crucial role in understanding public opinion and consumer behavior. In this paper, we conduct a comparative study of sentiment analysis on Amazon product reviews using two different sentiment analysis approaches: VADER (Valence Aware Dictionary and Sentiment Reasoner) and RoBERTa (Robustly Optimized BERT Approach). We explore the effectiveness of these models in capturing the sentiment expressed in Amazon reviews, considering various aspects such as accuracy, computational efficiency, and generalization capabilities. Our findings provide insights into the strengths and limitations of each approach, contributing to the advancement of sentiment analysis methodologies.

Keywords

Sentiment Analysis, Roberta, Vader, Machine Learning.

1. Introduction

Blog posts, online forums, social media, product review websites, etc. have become the primary means of expressing people's views and opinions since the age of the internet.

Millions of individuals use social networking sites like Twitter, Google Plus, Facebook, and others now a days to exchange ideas about their everyday lives and to express their opinions and feelings.

We have access to interactive media through online communities, where users may use forums to inform and influence others. A significant amount of sentiment-rich data is being produced via social media in the form of tweets, status updates, blog posts, comments, reviews, and other content.

The internet also gives businesses a chance to advertise by connecting them with consumers on a platform. When making decisions, people rely heavily on user-generated material found online. For example, before deciding to purchase a product or utilize a service, people usually check internet reviews and engage in social media discussions about it.

The amount of information generated everyday by users is very large for a normal human brain to digest. So, an automation technique known as sentiment analysis is widely used.

Sentiment analysis, sometimes referred to as opinion mining, is the act of locating and classifying text data that contains subjective information. As e-commerce sites like as Amazon, Flipkart, and Twitter have proliferated, it has become more crucial than ever for businesses to comprehend the opinions of their customers as reflected in their reviews in order to make wise judgements. Conventional techniques for sentiment analysis frequently depend on machine learning algorithms or lexicon-based strategies. Sentiment analysis is one area in which transformer-based models, such as

RoBERTa, have seen notable progress recently. In this study, we evaluate the effectiveness of a state-of-the-art transformer model, RoBERTa, against a rule-based sentiment analysis tool, VADER, on Amazon product evaluations.

1.1 Related Works

Previous research has explored various methodologies for sentiment analysis, including lexicon-based approaches, machine learning algorithms, and deep learning models. VADER, developed by Hutto and Gilbert, is a widely used lexicon-based approach known for its simplicity and effectiveness in capturing sentiment from social media texts. On the other hand, transformer models like BERT (Bidirectional Encoder Representations from Transformers) and its variants have achieved great performance in various natural language processing (NLP) tasks, including sentiment analysis. RoBERTa, an optimized version of BERT, has shown superior performance in understanding context and semantics in text data.

2. Sentiment Analysis

Opinion mining, another name for sentiment analysis, is a computer method for examining and deciphering the sentiment or emotional tenor of textual material. Sentiment analysis is now crucial for companies, organizations, and researchers because to the exponential rise in digital information created on social media, reviews, forums, and other online platforms.

Sentiment analysis's main objective is to automatically identify a text's sentiment polarity and categories it as positive, negative, or neutral. Various machine learning algorithms and natural language processing (NLP) approaches are used in this process to interpret the text's underlying attitudes, beliefs, and feelings.

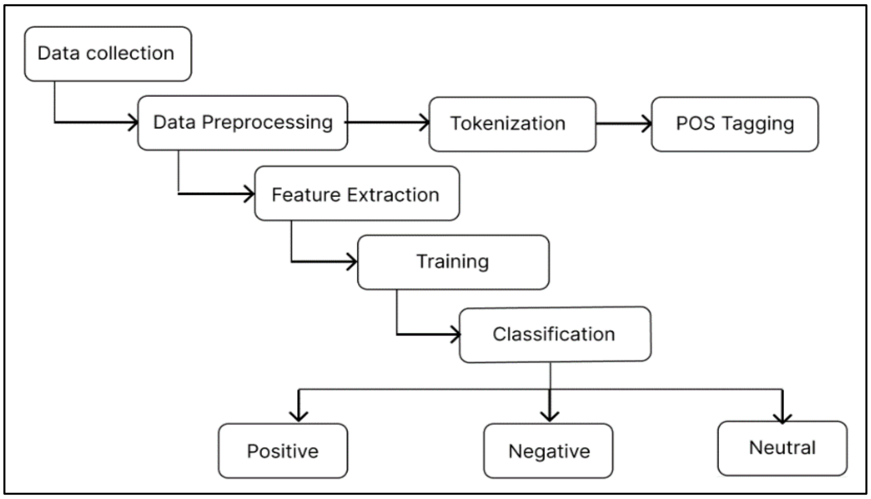
There are several approaches to sentiment analysis, including:

1. Lexicon-based Approaches: Sentiment lexicons or dictionaries with predetermined sentiment scores for words are the foundation of these techniques. How many and how strongly positive and negative words are used in a text determines its overall emotion. Veterans Aware Dictionary and Sentiment Reasoner, or VADER, and Sent WordNet are two examples of lexicon-based technologies.

2. Machine Learning Approaches: To categorize text into sentiment groups, machine learning techniques like naive Bayes, logistic regression, and support vector machines (SVM) are trained on labelled datasets. These algorithms use data-driven patterns and characteristics to forecast the emotion of unseen text.

3. Deep Learning Approaches: In sentiment analysis tasks, deep learning models—especially transformer-based architectures such as BERT (Bidirectional Encoder Representations from Transformers) and its variations, RoBERTa, and GPT—have demonstrated impressive performance. These models use fine-tuning methods and massive pre-trained language representations to extract context and intricate linguistic patterns from textual input.

Following are the phases or steps required for sentiment analysis of data:



2.1 Acquire data:

Sentiment analysis is applied to text data, which may need to be cleaned and processed. For example, parts of the text that convey no meaning, such as "the" or conjugations of a word, may need to be removed.

2.2 Data Processing

The processing of the data will depend on the kind of information it has. The processing of data is done to check for the unrelated data, missing values, or finding the errors in data if any. And thus such data is removed.

2.3 Feature Extraction

It is well-known to convert unprocessed data to numerical features that can be used while maintaining the information in the original data set. It can be completed automatically or manually. The data currently we are dealing in this paper is already available in number format.

2.4 Training

This is the phase where a model is trained with the available data. This training helps in classifying the new data in future. Also training helps the model to learn the type of data it is going to deal with.

2.5 Classification

Here the data is classified based on the desired parameters. For example, in our data we are classifying it into Positive, Negative or Neutral.
There can be multiple parameters.

3. Materials and Methods

Following python libraries are used to do the sentiment analysis of the data -

1. Punkt Sentence Tokenizer –

Using an unsupervised approach to create a model for words that start sentences, collocations, and abbreviations, this tokenizer splits a text into a list of sentences. Before it can be utilized, it needs to be trained on a sizable collection of target language plaintexts.

Sentence = "I have been drinking this tea for a long time now."

Tokens = ['I', 'have', 'been', 'drinking', 'this', 'tea', 'for', 'a', 'long', 'time', 'now', '.',]

2. Averaged perceptron tragger-

In NLTK, a pre-trained English part-of-speech (POS) tagger is called an averaged perceptron tagger. This is the model that loads with nltk.pos_tag. The tagger is in the averaged_perceptron_tagger.zip file. Labelling words in a sentence with the appropriate POS tags is known as POS tagging. The grammatical category of a word, such as noun, verb, adjective, adverb, etc., is indicated by POS tags.

sentence = "I have been drinking this tea for a long time now."

Result = [('I', 'PRP'), ('have', 'VBP'), ('been', 'VBN'), ('drinking', 'VBG'), ('this', 'DT'), ('tea', 'NN'), ('for', 'IN'), ('a', 'DT'), ('long', 'JJ'), ('time', 'NN'), ('now', 'RB'), ('.', '.'),]

3. Maxent ne chunker -

This named entity chunker in the Natural Language Toolkit (NLTK) for Python has already undergone training. It is a statistical model that locates and categories named entities—people, organizations, places, dates, and more—within a text by maximizing entropy. Its foundation is the ACE corpus, a sizable collection of text data with annotations.

```
Result = (S
I/PRP
have/VBP
been/VBN
drinking/VBG
this/DT
tea/NN
for/IN
a/DT
long/JJ
time/NN
now/RB
        ./.)
```

4. Results and Discussion

Our dataset was taken from the Amazon public domain. This labelled dataset was analyzed using a variety of feature extraction techniques. We employed a framework that applies a preprocessor to the raw sentences to improve their readability. Additionally, the dataset is trained with feature vectors by various machine learning approaches, and after that, semantic analysis provides a wide range of synonyms and similarity, which determines the polarity of the content.

1. Vader method

A natural language processing (NLP) tool called Valence Aware Dictionary for Sentiment Reasoning (VADER) determines the sentiment of a text by using a vocabulary of words and their semantic orientation. VADER is a rule-based analyzer that assesses the general sentiment of a sentence in addition to classifying words as positive, negative, or neutral. Based on a lexicon and rules that have been scientifically tested by numerous human judges, VADER is open-sourced under the MIT License.
Result of applying Vader method on data –

E.g. - Right now, I'm doing my work and I love to do my work.

Result = {'neg': 0.0, 'neu': 0.52, 'pos': 0.48, 'compound': 0.9487}.

Ran the Vader's method on entire dataset and classify the data as positive, negative or neutral.

Following image describes how the text is classified -

.

index	Id	neg	neu	pos	compound	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenominator	Score	Time	Summary	Text
0	1	0.0	0.695	0.305	0.9441	B001E4KFG0	A3SGXH7AUHU8GW	delmartian	1	1	5	1303862400	Good Quality Dog Food	I have bought several of the Vitality canned dog food products and have found them all to be of good quality. The product looks more like a stew than a processed meat and it smells better. My Labrador is finicky and she appreciates this product better than most.
1	2	0.138	0.862	0.0	-0.5664	B00813GRG4	A1D87F6ZCVE5NK	dll pa	0	0	1	1346976000	Not as Advertised	Product arrived labeled as Jumbo Salted Peanuts...the peanuts were actually small sized unsalted. Not sure if this was an error or if the vendor intended to represent the product as "Jumbo".
2	3	0.091	0.754	0.155	0.8265	B000LQOCH0	ABXLMWJIXXAIN	Natalia Corres "Natalia Corres"	1	1	4	1219017600	"Delight" says it all	This is a confection that has been around a few centuries. It is a light, pillowy citrus gelatin with nuts - in this case Filberts. And it is cut into tiny squares and then liberally coated with powdered sugar. And it is a tiny mouthful of heaven. Not too chewy, and very flavorful. I highly recommend this yummy treat. If you are familiar with the story of C.S. Lewis' "The Lion, The Witch, and The Wardrobe" - this is the treat that seduces Edmund into selling out his Brother and Sisters to the Witch.
3	4	0.0	1.0	0.0	0.0	B000UA0QIQ	A395BORC6FGVXV	Karl	3	3	2	1307923200	Cough Medicine	If you are looking for the secret ingredient in Robitussin I believe I have found it. I got this in addition to the Root Beer Extract I ordered (which was good) and made some cherry soda. The flavor is very medicinal.
4	5	0.0	0.552	0.448	0.9468	B006K2ZZ7K	A1UQRSCLF8GW1T	Michael D. Bigham "M. Wassir"	0	0	5	1350777600	Great taffy	Great taffy at a great price. There was a wide assortment of yummy taffy. Delivery was very quick. If your a taffy lover, this is a deal.

2. Roberta method –

Facebook AI researchers created the BERT (Bidirectional Encoder Representations from Transformers) model, which is the basis for the RoBERTa (short for "Robustly Optimized BERT Approach") model. Like BERT, RoBERTa is a transformer-based language model that processes input sequences and creates contextualized representations of words in a phrase by applying self-attention.

The fact that RoBERTa was trained on a larger dataset and with a more efficient training process is one of the main distinctions between it and BERT. Specifically, 160GB of text—more than ten times the size of the dataset used to train BERT—were used to train RoBERTa.

5. Comparison

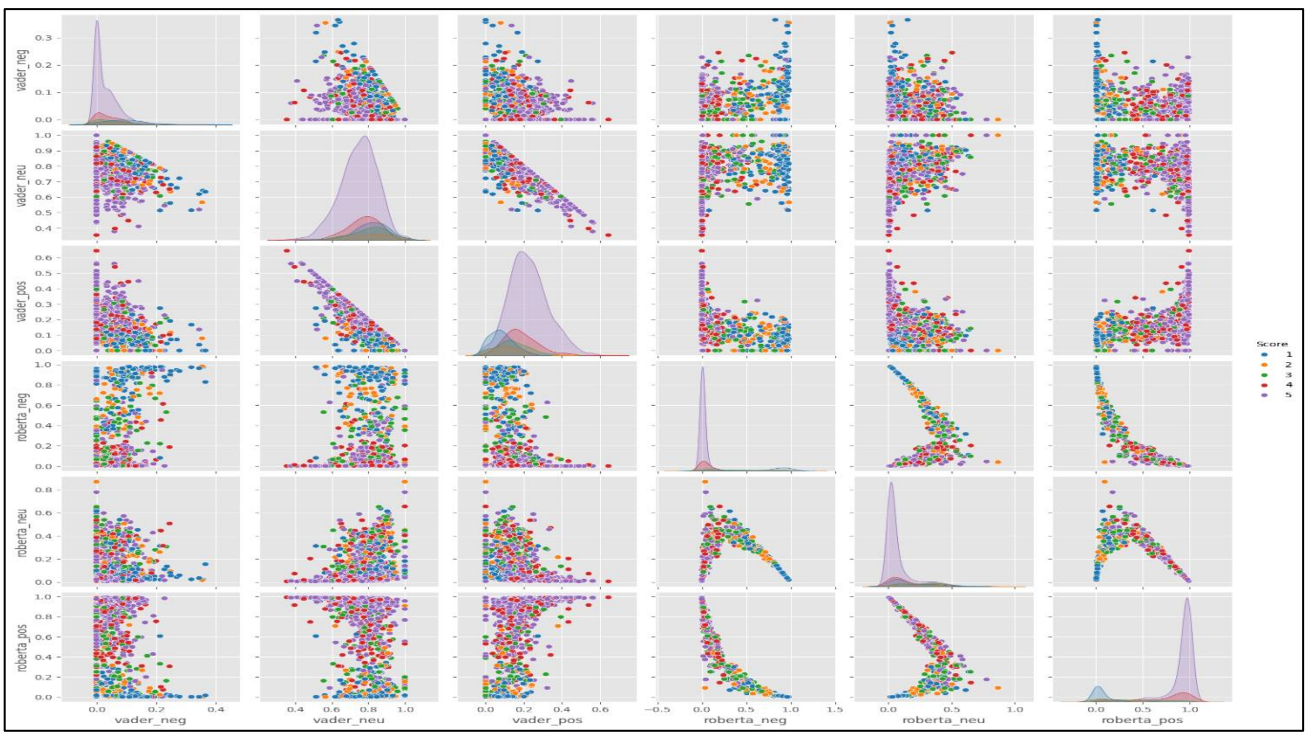
Comparison is done between Vader model and Roberta model by plotting a graph.

a. Accuracy and performance: When it comes to accuracy, Roberta usually beats Vader, especially when it comes to expressing complex emotions and expressions that depend on context. Roberta performs better because of its deep learning architecture, which allows it to learn from enormous volumes of data and adjust to different linguistic nuances.

b. Computational efficiency: Vader is suited for real-time or resource-constrained applications due to its computational efficiency. Roberta, on the other hand, is only useful in specific situations due to its deep learning architecture, which necessitates substantial processing power for training and inference.

c. Generalization: Comparing Roberta to Vader, who mostly depends on the sentiment lexicon and could have trouble with domain-specific language, Roberta shows stronger generalization ability across several domains and languages.

The following graph describes the comparison between Vader's method and Roberta's method :



6. Conslusion

Different benefits and trade-offs are presented by VADER and RoBERTa in sentiment analysis. RoBERTa outperforms VADER in terms of accuracy, context understanding, and generalization whereas VADER offers simplicity and efficiency. The exact needs of the application will determine which of the two approaches is best, considering elements like domain specificity, processing capacity, and accuracy.

7. Reference

[1] Samha,Xu,Xia, Wong & Li "Opinion Annotation in Online Chinese Product Reviews." In Proceedings of LREC conference, 2008.

[2]. L. Barbosa, J. Feng. "Robust Sentiment Detection on from Biased and Noisy Data". COLING 2010: Poster Volume. 36-44.

[3] P. D. Turney, "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews," in Proceedings of the 40th annual meeting on association for computational linguistics, pp. 417–424, Association for Computational Linguistics, 2002.

[4]Elli, Maria Soledad, and Yi-Fan Wang. "Amazon Reviews, business analytics with sentiment analysis." 2016

[5] Neethu M,S and Rajashree R," Sentiment Analysis in Twitter using Machine Learning Techniques" 4th ICCCNT 2013,at Ironhanded, India. IEEE – 31661

[6]Xu, Yun, Xinhua Wu, and Qin Xia Wang. "Sentiment Analysis of Yelp„s Ratings Based on Text Reviews." (2015).

[7] Rain, Callen. "Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning."Swarthmore College (2013).

[8] Bhatt, Aashutosh, et al. "Amazon Review Classification Sentiment Analysis." International Journal of Computer Science and Information Technologies 6.6 (2015): 5107-5110.

[9] J. Kamps, M. Marx, R. J. Mokken, and M. De Rijke, "Using wordnet to measure semantic orientations of adjectives," 2004.

[10]Chen, Weikang, Chihhung Lin, and Yi-Shu Tai."Text-Based Rating Predictions on Amazon Health & Personal Care Product Review." (2015)

[11]Shaikh, Tahura, and DeepaDeshpande. "Feature Selection Methods in Sentiment Analysis and Sentiment Classification of Amazon Product Reviews.",(2016)

[12]Nasr, Mona Mohamed, Essam Mohamed Shaaban, and Ahmed Mostafa Hafez. "Building Sentiment analysis Model using Graphlab." IJSER, 2017

[13] Pan S J, Ni X, Sun J T, et al. “Cross-domain sentiment classification viaspectral feature alignment”. Proceedings of the 19th international conference on World wide web. ACM, 2010: 751-760

[14]Text mining for yelp dataset challenge; Mingshan Wang; University of California San Diego, (2017)

[15] Elli, Maria Soledad, and Yi-Fan Wang. "Amazon Reviews, business analytics with sentiment analysis." 2016

[16] Xu, Kaiquan, et al. "Mining comparative opinions from customer reviews for Competitive Intelligence." Decision support systems 50.4 (2011): 743-754.

[17] Bollegala , D., Weir, D., & Carroll, J.. Cross-Domain Sentiment Classification using a Sentiment Sensitive Thesaurus. Knowledge and Data Engineering, IEEE Transactions on, 25(8), 1719-1731,2013

[18] Miao, Q., Li, Q., & Dai, R. (2009). AMAZING: A sentiment mining and retrieval system. Expert Systems with Applications, 36(3), 7192-7198.

[19] He, Ruining, and Julian ****McAuley. "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering." Proceedings of the 25th International Conference on World Wide Web.International World Wide Web Conferences Steering Committee, 2016

[20] V. M. K. Peddinti and P. Chintalapoodi, “Domain adaptation in sentiment analysis of twitter,” in Analyzing Microtext Workshop, AAAI, 2011.