Sentiment Analysis for Amazon Reviews to Identify Customer Interests.

K. Aloka Abeysirigunawardana
Department of Computer Systems Engineering
University of Kelaniya
aloka.abeysiri@outlook.com

Dr. Rasika Rajapaksha (Supervisor)
Department of Computer Systems Engineering
University of Kelaniya
rasikar@kln.ac.lk

Abstract— In the twenty-first century, consumers now frequently buy online and use reviews to judge the quality of the products they purchase. These reviews are also examined by businesses to enhance their offerings. In order to glean business insights from massive datasets, this study uses sentiment analysis on Amazon product reviews. To determine the specific subjects upon which the entire collection of reviews is predicated, an LDA model is created for the Amazon review dataset. Word frequencies for each topic are visualised by us. After determining which machine learning model is most appropriate for sentiment analysis, the analysis is carried out on a topic-by-topic basis. Using logistic regression, the themes derived from the LDA model are categorised as either positive or negative product review subjects. Analysis is done on both the positive and negative subjects.

Index Terms— Voice of Customer, Natural Language Processing, Information Retrieval, Machine Learning, Human-Computer Interaction, Bag of Words, Part of Speech, Natural Language Tool Kit, Latent Dirichlet Allocation

I. Introduction

Over the past ten years, online shopping has expanded significantly, surpassing all predictions and expectations. Customers in online shopping systems are more likely to rely on reviews of the goods they plan to purchase than on word-of-mouth recommendations [1]. IGI Global states that a consumer's online review is an assessment of a particular online product or service that they have used [2]. These reviews are utilised for more than just that, though; they are also employed to perform statistical analysis to determine consumer behaviour following the purchase of a good or service, which aids in the creation of corporate plans for expansion[3].

Online reviews are important, according to statistics. 94% of internet shoppers check reviews before making a purchase, per Fan & Fuel (2016). 95% of consumers check reviews before making a purchase, according to Spiegel Research Centre (2017), particularly when it comes to product-specific information. [3] [4].

Voice of the Customer (VoC) programs, customer service experiences, social media sentiment analysis, product experiences, brand sentiment analysis, market research, and more are examples of these commercial uses [5]. It is challenging to comprehend a product's client base, market reach, audience kind, customer behaviour, and hidden interests—all of which are important product insights—by simply reading reviews.

Analyse the data lake using machine learning techniques to gain profound insights into the products [8]. The significance of customer reviews as a source of data to delve
deeply into consumer opinions and their future interests is emphasised by this study. The study is carried out in two ways. Initially, a thorough statistical examination of a publicly accessible dataset comprising different Amazon Alexa reviews. The dataset is available at the following citation [9]. Secondly, a sentiment analysis allowing large-scale processing of data to these Amazon reviews in an efficient and cost-effective manner [12].

Latent Dirichlet Allocation (LDA) was selected as the topic modeling technique due to its effectiveness in identifying underlying thematic structures within large, unstructured datasets like customer reviews. LDA is particularly valuable for this study as it offers interpretability and helps uncover distinct topics, allowing businesses to better understand common customer concerns and preferences. Prior to applying the LDA model, several preprocessing steps were conducted to ensure data quality and relevance. This included removing noise (such as irrelevant words or special characters), handling duplicates to prevent data skewing, and filtering out reviews that did not provide meaningful feedback. These steps enhance the rigor of the analysis by ensuring that the dataset used is clean and representative, ultimately providing more accurate insights from the sentiment analysis [28].

II. Related work

Any opinion, attitude, or judgement brought on by a sensation is referred to as sentiment. Sentiment analysis is essentially the study of human sentiment on a topic; to put it simply, it is the attitude towards anything [10][11]. Opinion analysis is another name for sentiment analysis. The widespread use of the Internet has made it quite simple for people to voice their opinions. Users and viewers can fill out forms and forums, review, discuss, and offer feedback on the things they have purchased, as well as on various social networking platforms. Sentiment analysis has adequately established the basics thanks to these enormous data collections made possible by web platforms. However, the sentiment analysis procedure now has a number of shortcomings as a result of these online data. The first fault found is that the online text collected would not be highly guaranteed because anyone can publish information on the internet. While some reviews may be taken into consideration, some might be regarded as phoney viewpoints. The second noted problem is that the actual classification, sometimes referred to as the ground truth, is not included with the data when it is given. This shows if the review is actually favourable or unfavourable [12]. In order to identify and extract information about the subjectivity of the material, sentiment analysis draws on a variety of technological issues, including text analysis, natural language processing, and computational linguistics [13]. Sentiment analysis is widely used in modern industries for everything from corporate marketing to customer service. Finding the writer's viewpoint and opinion and assessing such parameters in light of one or more issue domains are the major objectives of sentiment analysis. The entity level, word level, phrase level, and document level are the four primary forms of sentiment analysis that are now carried out [14].Numerous technical ideas, including machine learning (ML) methods and natural language processing (NLP), are included in sentiment analysis. Simply enough, sentiment analysis determines if a message is neutral, positive, or negative by training a model on millions of text fragments [15].By using this method, we create a lexicon—a unit of a language—that contains information on which words, phrases, or expressions have positive meanings and which have negative ones. For instance, each WordNet uses SentiWordNet as a lexical resource. For every language, Synset offers three numerical scores: positive, negative, and objective [16].

Fig. 1: Sentiment analysis workflow

The number of people who shop and buy online has increased significantly over the last two years due to the COVID-19 pandemic, reaching over 100%. In a similar vein, the proportion of people who use the Internet has increased rapidly and is expected to continue doing so [18]. According to a 2011 survey, 74% of consumers trust the suggestions; however, in 2012 and 2013, that number fell to 60% and 57%, respectively. In 2019, however, the findings have shifted, indicating that up to 94% more consumers now trust online sentiment reviews [19].

To uncover hidden subjects in customer evaluations, topic modelling techniques such as Latent Dirichlet Allocation (LDA) have been used in addition to sentiment analysis. Text corpora classification and organization have benefited greatly from LDA. While sentiment analysis has made significant strides, challenges such as detecting sarcasm and processing domain-specific language remain areas of active research [18]. In addition to LDA, other topic modeling techniques, such as Non-Negative Matrix Factorization (NMF) and BERT-based models, have been explored in similar studies for organizing and classifying textual data. While methods like BERT are effective for capturing contextual relationships in text, LDA remains advantageous in its interpretability and efficiency for identifying overarching themes within large datasets, making it suitable for this study. Furthermore, ethical considerations such as data anonymization and compliance with data privacy standards are crucial when handling customer review data. These measures are essential to protect individual privacy and ensure that sentiment analysis is conducted responsibly, especially as insights from automated analyses increasingly inform business decisions [28].

III. Methodology

The purpose of this study is to generate business insights by analysing Amazon product evaluations using topic modelling and sentiment analysis techniques. Data preparation, topic modelling using Latent Dirichlet Allocation (LDA), and sentiment categorisation using machine learning models were all part of the methodical approach. This section outlines the methodology used, detailing each step in the process, including data handling, modeling, evaluation metrics, and final classification [28].

A. Dataset

The dataset consists of 3150 reviews from Amazon consumers who bought the Amazon Alexa product between May and July of 2018 make up the dataset [7]. Numerous studies have been done on the publicly accessible dataset, which can be found at www.kaggle.com. Five columns of data are included in the Amazon Alexa customer data: "rating" (comment score), "date" (comment date), "variation" (16 Alexa product variants), "verified_reviews" (the comment), and "feedback" (binary sentiment score based on the rating). The data includes Alexa product reviews for three months.

Fig. 2: Header of the dataset

	ratin
0	
1	
2	
3	
4	

B. Data Exploration and Preprocessing

In order to comprehend review trends and consumer feedback, the technique started with an analysis of the dataset, which involved creating descriptive statistics, analysing rating distributions, and summarising its structure. To find trends, data visualisation using histograms was also done. The text was then prepared for analysis by data preparation, which included tokenisation, stemming, lemmatisation, stopword removal, and vectorisation using the Bag of Words model [20], [21]. Creating the corpus—the body of tokenised material needed for topic modeling—was the next stage. A document-term matrix (DTM) was created at this phase, in which the frequency of words found in each document (review) is used to represent it. Furthermore, a dictionary known as id2word was developed, which gives every word in the corpus a distinct identifier. The corpus and id2word dictionary serve as the two main inputs for the LDA model. The document-term matrix provides the frequency of words in each review, while the dictionary enables the model to map each word to its corresponding identifier [28].

C. Topic Modeling using LDA

The Amazon reviews' latent thematic structures were found using the Latent Dirichlet Allocation (LDA) model. The generative probabilistic model known as LDA makes the assumption that every review is a combination of multiple themes, each of which is represented by a word mixture. Finding the underlying consumer interests and concerns in the product reviews was the goal. Every topic is represented as a distribution over words, and every word in the document is given a probability distribution over a predetermined number of topics [22].One of the most important steps in LDA is determining the ideal number of subjects (k). To assess how effectively the themes reflected the underlying structure of the reviews, many models with different k values were developed. Two criteria were used to assess these models' performance: topic coherence and perplexity [23]. Because Latent Dirichlet Allocation (LDA) is so good at finding underlying subject structures in big, unstructured datasets like customer reviews, it was chosen as the topic modelling technique. For this study, LDA is very useful since it provides interpretability and helps identify unique subjects, which enables organisations to better grasp typical client preferences and concerns. To guarantee the quality and relevance of the data, a number of preprocessing procedures were carried out before the LDA model was used. This involved filtering, handling duplicates to avoid data skewing, and eliminating noise (such as superfluous words or special characters) reviews that did not provide meaningful feedback [28].

- **Perplexity** measures how well a probabilistic model predicts a sample. In topic modeling, it evaluates how surprised the model is when presented with unseen data. Lower perplexity values indicate that the model is better at capturing patterns in the data. However, perplexity alone is not always a reliable metric for evaluating topic models, as it tends to favour complex models.[23].
- **Topic Coherence** is another key evaluation metric that measures the semantic similarity between high-scoring words within a topic. Unlike perplexity, coherence is a human-centered measure that reflects how interpretable and meaningful the topics are to humans. A high coherence score indicates that the top words in a topic are closely related and make sense as a collective theme. For this study, the coherence score was prioritized as the primary metric, as it is a better indicator of meaningful and interpretable topics. The coherence score obtained for the optimal model was 0.476, which indicated a reasonably good level of interpretability [23].

D. Model Evaluation and Visualization

Following the training and evaluation of the LDA model, the topics were visualised using the pyLDAvis package, a potent tool for interpreting LDA results. The visualisation offers an inter-topic distance map, in which each circle represents a topic and its size indicates the percentage of reviews attributed to that topic; larger circles indicate more frequent topics, and the distance between circles indicates how distinct the topics are from one another [24]. Additionally, the pyLDAvis tool enables the visualisation of the most pertinent words for each topic, which aids in comprehending the thematic structure of the reviews by exposing customer sentiments, product problems, or other underlying concerns [25].

E. Sentiment Classification using Machine Learning

Following topic identification, a variety of machine learning models were trained to categorise the topics into positive or negative feelings based on the feedback column. These models included Logistic Regression, KNN, SVM, Bagged Decision Trees, Random Forest, and ANN [26]. Accuracy metrics were used to assess each model's predicted performance. The model that performed the best was logistic regression, which had the most accuracy in determining whether a particular review was positive or negative.

F. K-Fold Cross Validation

Each machine learning model's performance was assessed using cross-validation. In particular, the generalisability of the model was evaluated using k-fold cross-validation (k=5 and k=10). This approach divides the dataset into k divisions and randomly shuffles it. The model is tested on the remaining partition after being trained on k-1 partitions for each partition. Model performance is evaluated by averaging the assessment score from each fold [27].

G. Prediction of New Reviews

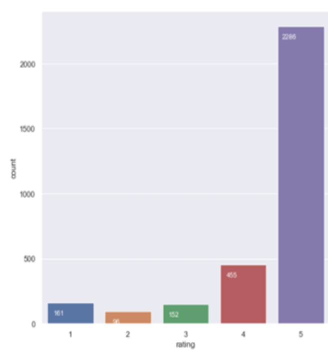
The sentiment of fresh, unseen reviews was predicted using the learnt logistic regression model. To forecast its sentiment, the new review was tokenised, converted into a numerical feature vector, and then run through a logistic regression model. The label with the highest probability was chosen as the projected mood after the model produced a probability distribution over both positive and negative labels [28].

IV. Results

A. Exploratory Data Analysis on Dataset

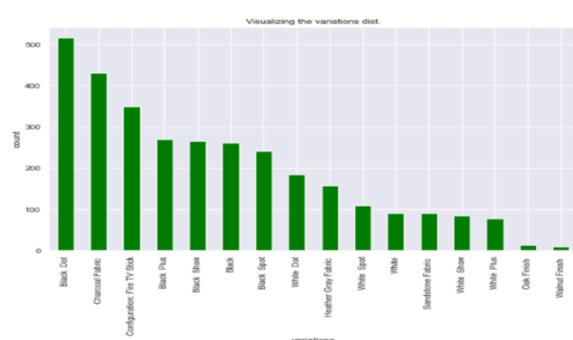
Python programs are used for the initial analysis of the dataset. Column data types are revealed by the data.info() function. To evaluate the sentiment distribution, reviews are grouped by rating (1–5) and displayed. To guarantee thorough depiction of product lines, feedback and the variety of 16 Alexa device types are also analysed [28].

Fig. 3: Rating and feedback distributions.



The distribution of the dataset is based on the range of Alexa devices. The dataset contains 16 distinct variants, with the distribution of those varieties as follows.

Fig. 4: Verity of devices with the count of reviews Distribution.



The most common positive and negative terms in the reviews are highlighted in a word cloud. Insights into recurring topics and user sentiments are provided by this visualisation, which will be helpful for further sentiment analysis.

Fig. 5: Word Cloud Positive and Negative

Task 01: Clustering Review and Identify Topics

B. Measuring the Model

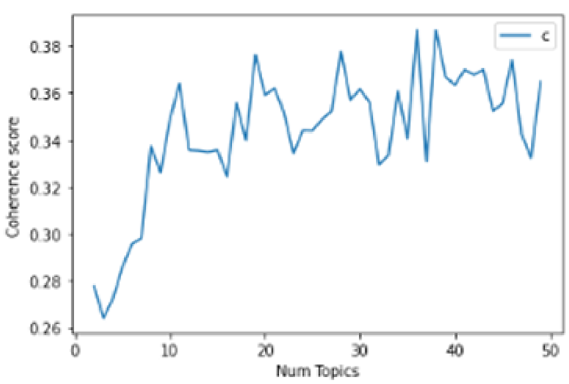
Both coherence and perplexity ratings are computed to evaluate the LDA model. For example, the following perplexity and coherence values are achieved when the topic count is set to 20: Coherence Score: 0.3185, Perplexity: -10.3730

These metrics are used as standards to assess the usefulness and pertinence of the subjects that the model identified [28].

C. Finding the Best Model

The coherence values for various topic counts (k) are used to calculate the ideal number of subjects. To help choose the best model for additional research, the coherence score is plotted versus the number of themes [28].

Fig. 6: Coherence value against number of topics.



There is also a table that summarises the coherence values for topic counts between 2 and 49. When there are 40 themes and a coherence score of 0.4041, the ideal model is found.

TABLE I: Coherence Value for Topics 2 - 49.

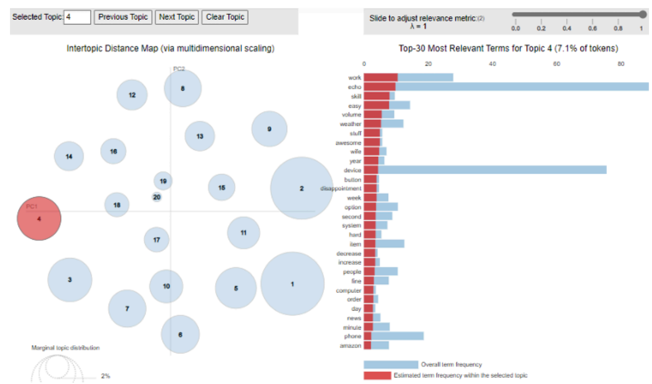
Num Topics = 2 has Coherence Value of 0.2288
Num Topics = 3 has Coherence Value of 0.2843
Num Topics = 4 has Coherence Value of 0.2834
Num Topics = 5 has Coherence Value of 0.3151
Num Topics = 6 has Coherence Value of 0.3377
Num Topics = 7 has Coherence Value of 0.2885
Num Topics = 8 has Coherence Value of 0.335
Num Topics = 9 has Coherence Value of 0.3219
….
Num Topics = 45 has Coherence Value of 0.3699
Num Topics = 46 has Coherence Value of 0.3487
Num Topics = 47 has Coherence Value of 0.3812
Num Topics = 48 has Coherence Value of 0.3708
Num Topics = 49 has Coherence Value of 0.359
38 : Optimal Topic number = 40 Coherence Scores = 0.4041

D. Visualization for 20 Topics

To give a thorough understanding of how various topics are represented in the dataset, a visualisation of the LDA model with 20 themes is displayed. The thematic organisation of the evaluations is highlighted by displaying each topic with its most pertinent phrases [28].

Fig. 7: Visualization for 20 topics.

Topic 04

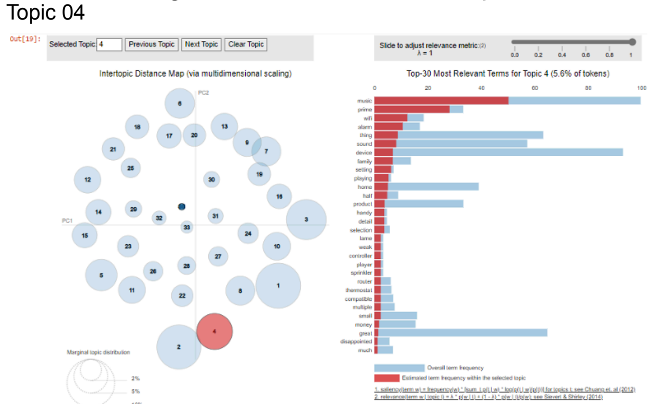


E. Visualization for 40 Topics – Best Fit LDA Model

The 40-topic model, which was determined to be the best fit based on the coherence score, produces a similar visualisation.

Fig. 8: Visualization for 40 topics.

Topic 04



Task 02: Comparison of Machine Learning Models Used to Identify the Type of Review

The review kinds are categorised using a variety of machine learning algorithms. To determine which model performs best, the accuracies of each model—obtained using both k-fold cross-validation (k = 5 and k = 10) and normal test-train splits—are compared [28].

TABLE 2: ML Best Model Accuracies.

Created Model Name	Normal test and train split of 33% and 67% respectively	K fold cross validation k=5	K fold cross validation k=10
Logistic Regression	0.949	0.946	0.948
K-Nearest Neighbors (KNN)	0.886	0.933	0.937
Support Vector Machine (SVM)	0.930	0.923	0.925
Bagged Decision Trees (Bagging)	0.934	0.929	0.932
Random Forest	0.943	0.934	0.937
Neural Network	0.924	0.918	0.918

F. Topic Tagging

The identified keywords are used to tag every topic that the Logistic Regression model predicts. With both favourable and negative reviews labelled appropriately, these subjects offer insights into the character of the reviews.

TABLE 3: Topic Tagging Results.

Topic	Tag
Topic: setup alarm many easy screen alexa several install simple home	A Positive Review
Topic: alexa part bedside ability instruction search clock music computer variety	A Positive Review
Topic: account device bluetooth echo manual personal membership respond star system	A Negative Review
Topic: equipment connectivity kid well good interact fast song dictionary player access need	A Positive Review
Topic: item purchase next work blue ready connect button echo light	A Positive Review
Topic: device echo home wifi problem original night well audio thing	A Positive Review
Topic: move card face dumb spot stupid super wall star clock	A Negative Review
Topic: house clock camera basic good show time display thing home	A Positive Review
Topic: volume enable speaker bass look less overdriven try34 time screen	A Positive Review
Topic: future bass exact surprised time fact small voice money well	A Positive Review
Topic: music love alarm light feature thing youtube great weather good	A Positive Review
Topic: thermostat music product tablet phone disappointed application	A Positive Review
Topic: phone number service device access able kind amazon sprint minimum	A Positive Review
Topic: equipment connectivity kid well good interact fast song dictionary player access need	A Positive Review
Topic: wifi speech router recognition office area thing wife intercom reception	A Positive Review
Topic: music sound quality prime speaker home need family device product	A Positive Review
Topic: gift birthday order package large step porch the open damage says but	A Positive Review
Topic: equipment connectivity kid well good interact fast song dictionary	A Positive Review
Topic: product nice star device loud work reason sure complaint alone	A Positive Review
Topic: device answer word echo little people screen time replacement light	A Positive Review
Topic: free screen month useful life bulb video device alexa good	A Positive Review
Topic: able price stupid music one thermostat rock dropping question else	A Positive Review
Topic: echo good sound day voice function weather convenient speaker	A Positive Review
Topic: great speaker happy weird reason time music noise interact major	A Positive Review
Topic : first fire amazon stick device horrible warranty againthis year major	A Positive Review
Topic : room program living request thing sling time respond speaker first	A Positive Review
Topic : trouble fault alexa apple product amazon speaker time music none	A Negative Review

V. Discussion

The dataset, which includes 3,150 reviews of different Amazon Alexa products, is overwhelmingly positive, with more than 87% of the evaluations receiving favourable ratings. A bias in the dataset is reflected in the large percentage of positive evaluations, which made it difficult to classify sentiment in a fair manner. According to the rating distribution shown in the preceding chapter, most reviews received ratings of five or four stars, with very few receiving ratings lower than three. Even though there are fewer negative reviews, they are essential for pointing out possible areas where the product could be improved, but it is more difficult to glean meaningful insights from them due to the imbalance in sentiment [9 Word clouds created for both positive and negative evaluations' textual content draw attention to various consumer problems. Words like "device," "still," and "time" predominate in unfavourable evaluations, whereas "great," "love," and "music" are often in positive ones. According to these linguistic patterns, unfavourable reviews highlight certain problems like functionality and performance, but good evaluations frequently highlight the user experience and emotional bond with the product. [26][28].

A. **Topic Identification Using LDA**

Providing insights into customer reviews by grouping them into different areas was one of the main goals of this study. To make it easier to read consumer feedback, the LDA model was used to aggregate related reviews into cohesive subjects. A 40-topic model was found to have the best match after 20 topics were first developed and their coherence and perplexity scores were evaluated. The ideal number of subjects to capture the variety of customer feedback while preserving interpretability was determined via the topic coherence measure [25].

A strong foundation for spotting trends and patterns in the dataset was provided by the 40-topic LDA model. For instance, it assisted in classifying consumer grievances regarding "Bluetooth connectivity" and "speaker quality," while simultaneously highlighting highly acclaimed aspects like "voice recognition" and "ease of setup." A structured overview of client sentiments was provided by classifying each of these subjects as either favourable or negative. This procedure showed how topic modelling might transform unstructured data into insights that could be put to use. [28].

G. **Performance of Machine Learning Models**

Following the identification of topics, machine learning models were employed to classify these topics as positive or negative. The models used for this task included Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Trees, Random Forest, and a neural network. Logistic Regression outperformed the other models in terms of accuracy, making it the most suitable choice for predicting feedback polarity in this context. While the dataset imbalance posed challenges for some models, cross-validation helped mitigate overfitting and provided a clearer understanding of model performance [28].

H. **Business Insights from Topic Tagging**

By classifying the themes produced by the LDA model as either positive or negative, the primary objective of this study was to derive business insights from the customer reviews. The examination of the 40 different subjects found 37 favourable items and 3 negative ones. The subjects found were nonetheless helpful for comprehending customer wants and possible product changes, even though the disparity between positive and negative responses reduced the depth of the negative insights. Negative subjects like "speaker faults" and "Bluetooth connection issues" gave Amazon particular places to concentrate on improving its products. Positively, discussions about "ease of setup" and "functionality for kids" brought attention to aspects of the product that consumers value. By examining these subjects, the study provides insightful information that can directproduct development and customer service strategies [28].

I. **Limitations**

There were a number of restrictions on this study. The size and imbalance of the dataset, which impacted the conclusions' generalisability, were the most important of these. The models trained on this data might not perform as well on more balanced datasets because it only has 3,150 reviews and has a significant bias towards positive remarks. Another drawback was the incapacity to identify more complex expressions, such irony and sarcasm, which have a big impact on sentiment categorisation accuracy. Furthermore, even though topic modelling offered insightful information, some themes were difficult to understand because of the ambiguity of some terminology.

Last but not least, this study did not thoroughly investigate automatic topic labelling, which restricted the approach's scalability. By investigating bigger, more balanced datasets and applying more sophisticated methods for autonomous topic labelling, future studies could overcome these constraints [28].

VI. CONCLUSION

In order to extract useful business insights, this study successfully performed sentiment analysis on Amazon Alexa product reviews. In order to obtain a thorough grasp of consumer feedback, the study started with a thorough examination of the dataset using visual aids. The reviews were successfully prepared for analysis through the use of data preparation techniques. 40 coherent topics were found after a Latent Dirichlet Allocation (LDA) model was developed and optimised. These topics were then classified as positive or negative using a variety of machine learning models, such as neural networks, K-Nearest Neighbours (KNN), Support Vector Machines (SVM), Decision Trees, Random Forest, and Logistic Regression. Interestingly, the model that performed the best for sentiment categorisation was Logistic Regression, proving its usefulness in this situation. The research's conclusions offer insightful information on consumer sentiment, pointing out both areas of satisfaction and possible enhancements for the product. By identifying 37 positive and 3 negative subjects, Amazon is able to solve difficulties like "Bluetooth connection problems" and "speaker faults" while also identifying qualities that appeal to buyers, including "ease of setup" and "functionality for kids." In addition to improving customer happiness, this dual focus helps guide marketing and product development [28].

These observations can help Amazon improve customer service and streamline their product selection. While resolving negative feedback can result in focused adjustments that boost user happiness and loyalty, the company can further highlight the identified favourable traits in marketing initiatives. Additionally, the study emphasises how crucial it is to continuously monitor consumer feedback using sentiment analysis, which helps firms stay flexible and adapt to the ever-changing needs of their clientele The study points out that future work could benefit from automatic labelling approaches, especially as the number of data expands, even if manual labelling of subjects was doable due to the small topic collection. Important areas for future research include addressing issues like irony and sarcasm in sentiment analysis and enhancing model performance on unbalanced datasets. Future research can offer even more detailed insights into customer attitude and preferences by improving these areas, which will ultimately lead to more successful company strategies [28].

VII. REFERENCES

[1] Pattern, "Why Reviews Are Important on Amazon," Pattern.https://pattern.com/blog/why-are-reviews-on-amazon-important/ (accessed Jun. 20, 2022).

[2] Publons, "Publons.com." https://publons.com/ (accessed Jun. 20, 2022).

[3] Revain, "Why Are Customer Reviews So Important?," Medium.https://medium.com/revain/why-are-customer-reviews-so-important-185b915d4e5d (accessed Jun. 20, 2022).
[3] Revain, "Why Are Customer Reviews So Important?," Medium.https://medium.com/revain/why-are-customer-reviews-so-important-185b915d4e5d (accessed Jun. 20, 2022).
[4] Simplify Everything, "8 statistics on why you should care about local SEO," https://simplify-everything.com/en/blog/2019/10/17/8-statistics-on-why-you-should-care-about-local-seo-infographic/ (accessed Jun. 20, 2022).
[5] S. Yi and X. Liu, "Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review," Complex Intell. Syst., vol. 6, no. 3, pp. 621–634, Oct. 2020, doi: 10.1007/s40747-020-00155-2.
[6] AMZ Insight, "Importance of Amazon Product Review and Rating for Sellers," Dec. 08, 2016. https://amzinsight.com/amazon-product-review-importance/ (accessed Jun. 20, 2022).
[7] Kaggle, "Sentiment Analysis of Alexa," https://kaggle.com/khushi08/sentiment-analysis-of-alexa (accessed Jun. 20, 2022).
[8] H. Zhang, Sentiment Analysis on Amazon Reviews, 2019, doi: 10.13140/RG.2.2.31090.53447.
[9] E. Gokce, "Sentiment Analysis on Amazon Reviews," Towards Data Science. https://towardsdatascience.com/sentiment-analysis-on-amazon-reviews-45cd169447ac (accessed Jun. 20, 2022).
[10] X. Fang and J. Zhan, "Sentiment analysis using product review data," J. Big Data, vol. 2, no. 1, p. 5, Jun. 2015, doi: 10.1186/s40537-015-0015-2.
[11] S.-M. Kim and E. Hovy, "Determining the Sentiment of Opinions," in COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics, Geneva, Switzerland, Aug. 2004, pp. 1367–1373. Accessed: Jun. 20, 2022. [Online]. Available: https://aclanthology.org/C04-1200.
[12] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," Found. Trends Inf. Retr., vol. 2, no. 1–2, pp. 1–135, Jul. 2008, doi: 10.1561/1500000011.
[13] P. D. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," arXiv, Dec. 11, 2002. Accessed: Jun. 22, 2022. [Online]. Available: http://arxiv.org/abs/cs/0212032.
[14] Keyhole, "How Does Sentiment Analysis Work?," Keyhole, Aug. 29, 2020. https://keyhole.co/blog/how-does-sentiment-analysis-work/ (accessed Jun. 22, 2022).
[15] G. Jeh and J. Widom, "Mining the Space of Graph Properties," Mar. 2003, doi: 10.1145/1014052.1014075.
[16] B. Liu, Sentiment Analysis and Subjectivity, p. 38.
[17] R. Brown, "Sentiment Analysis: How it Works, Types & Everything You Need to Know," Medium, Mar. 04, 2021. https://cogitotech.medium.com/sentiment-analysis-how-it-works-types-everything-you-need-to-know-822a1f2ddeaf (accessed Jun. 22, 2022).
[18] D. Mohey El-Din, "Analyzing Scientific Papers Based on Sentiment Analysis (First Draft)," 2016, doi: 10.13140/RG.2.1.2222.6328.
[19] W. Karwowski, Advances in the Human Side of Service Engineering, Jan. 2012. Accessed: Jun. 24, 2022. [Online].
[20] "GreTap94.pdf." Accessed: Jun. 27, 2022. [Online]. Available: in.net/courses/07/nlp/Gre:Tap:94.pdf.
[21] D. D. Palmer, "Chapter 2: Tokenisation and Sentence Segmentation," p. 23.
[22] "Perplexity To Evaluate Topic Models." http://qpleple.com/perplexity-to-evaluate-topic-models/ (accessed Jun. 29, 2022).
[23] R. Kulshrestha, "A Beginner's Guide to Latent Dirichlet Allocation (LDA)," Towards Data Science. https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2 (accessed Jun. 27, 2022).
[24] R. Kulshrestha, "Latent Dirichlet Allocation (LDA)," Medium, Sep. 28, 2020. https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2 (accessed Jun. 27, 2022).
[25] Analytics Vidhya, "Topic Modeling and Latent Dirichlet Allocation (LDA) using Gensim," Analytics Vidhya, Jun. 28, 2021. https://www.analyticsvidhya.com/blog/2021/06/part-2-topic-modeling-and-latent-dirichlet-allocation-lda-using-gensim-and-sklearn/ (accessed Jun. 27, 2022).
[26] K. Ganesan, "Build Your First Text Classifier in Python with Logistic Regression," Kavita Ganesan, PhD, Jul. 25, 2019. https://kavita-ganesan.com/news-classifier-with-logistic-regression-in-python/ (accessed Jun. 25, 2022).
[27] J. Brownlee, “A Gentle Introduction to k-fold Cross-Validation,” Machine Learning Mastery, May 22, 2018. https://machinelearningmastery.com/k-fold-cross-validation/ (accessed Jul. 02, 2022).
[28] K. A., A. (2022). Sentiment Analysis for Amazon Reviews to Identify Customer Interests [Thesis of Research]. University of Kelaniya.