MSc RESEARCH PROJECT: 771764_B23_T1

ANALYSIS OF THE REPUTATION OF AMAZON USING SENTIMENT ANALYSIS

ABSTRACT

Sentiment analysis, often known as opinion mining, is a natural language processing (NLP) technique that determines whether textual data is neutral, positive, or negative. It is impossible to underestimate the significance of consumer opinions regarding a product or service since they have the power to influence a company's sales, which determines whether it succeeds or fails in the industry. The objective of the research is to examine how sentiment analysis can be used to provide accurate and comprehensive assessment of a company's reputation based on customers' reviews. This study was carried out using Amazon shoes reviews datasets consisting of customers' reviews and their ratings. The datasets were cleaned, preprocessed and the features of the dataset was extracted using bag of words(CountVectorizer) and term frequency - inverse document frequency(TF-IDF). Random forest and support vector machine(SVM) models were explored in the machine learning algorithms on the test datasets. Furthermore, the deep learning algorithms for natural language processing, specifically focusing on DistilBERT and Long Short-Term Memory (LSTM) were explored to train the amazon shoes reviews corpus.

The outcome of this study shows that in the Machine learning models, random forest for BOW and TF-IDF produced accuracy of 86% and 87% while SVM also produced accuracy of 86% and 85% for BOW and TF-IDF respectively on the test datasets. Deep learning algorithms produced high performance of accuracy of 90% for LSTM model and 80% for DistilBert model when used on the train datasets of the Amazon shoes reviews. As a vital information tool, sentiment analysis assists organizations to optimize their products and services. by providing objective customer insights, analyze large customers' data and generate real time results

INTRODUCTION & BACKGROUND

E-commerce like Amazon allow users to buy conveniently, save time, and have their purchases delivered to their locations at the earliest time possible.

Kovalova et al (2021) in their research proposed that argued that a company's reputation as an intangible asset, is capitalized and constitutes a significant component of its worth. It also serves as the foundation for enhancing the company's sustainability, growth, and financial performance, boosting sales and revenue.

A wide range of elements frequently influence the decision-making process when consumers shop online. A critical component is the assessment of product reviews, a procedure with clues that can either increase or decrease a product's perceived legitimacy.

Hong & Pittman (2020). through the analysis of three crucial cues—star ratings, the quantity of reviews, and the valence of those reviews—explored the dynamics of online reviews on a platform that is similar to Amazon in an effort to better understand the nuances of customer decision-making. The discovered that the quantity of reviews is a good indicator of the strength of the argument, with negatively rated reviews typically outweighing star ratings. Participants believed a high star ranking indicated greater credibility and believed the system when a review was positive. But in the case of a bad review, participants base their judgement more highly on a large number of reviewers.

Kim & Kim (2022). Today's customers' base their purchasing decisions on the information they may find on internet shopping platforms. This information posted on online platforms can have a substantial impact on their purchase behavior. They investigated the Latent Dirichlet Allocation (LDA) topic modelling and found that, contrary to what might be expected, star ratings do not necessarily translate into higher sales. Instead, they appear to have an inverse U-shaped association. Nonetheless, the findings show that the volume of star ratings has a favourable impact on sales. According to their research, reviews that focus on product quality and added value increase sales.

Sharma, Chakraborti, and Jha (2019) conducted research to determine how Amazon book sales are influenced by online reviews. The primary objective was to delve into the intricate relationship between online reviews and book sales on the Amazon platform. This research investigates the efficacy of modelling techniques, such as regression analysis, decision-tree analysis, and artificial neural networks, for forecasting book sales by applying a range of pertinent attributes and their interactions as predictive variables at amazon.in. Online reviews' polarity is measured by sentiment analysis. The outcome demonstrates that the decision-tree based model and regression analysis are outperformed by the artificial neural network model. The research revealed that online reviews held a position of paramount importance in the decision-making process of Amazon customers. It became evident that these reviews had a profound impact on the perceptions and expectations of consumers, thereby significantly influencing their purchasing decisions and even product pricing dynamics in the highly competitive online marketplace.

In recent years, machine learning techniques have become increasingly prevalent in sentiment analysis due to the ongoing advancements in artificial intelligence. The most popular machine learning algorithms among them are Bert, Naive Bayes and Support Vector Machines.

Jihua Cao et al. (2023) employed text mining and deep learning techniques to perform sentiment analysis on user-generated online reviews from Taobao and JD.com, two e-commerce platforms. The latent Dirichlet allocation (LDA) deep learning model and the long short-term memory (LSTM) deep learning tool, Word2Vec for word vector training are used to develop the Consumer online reviews–Extract short text–Sentiment analysis–Cluster feature (CESC) methodology for improving product. Based on customer online reviews, the model can accurately determine the attributes and attitudes of the products that customers find desirable. Compared to classical machine learning, the accuracy of the model is 89.52%, the LSTM network appears to have an impressive learning impact on the processing of texts.

Transformer based models have been seen to generate great results in sentiment analysis and many Natural Language tasks in recent times. Chouikhi & Jarray (2023) used ensemble BERT models (Arabic BERT, mBERT, and AraBERT) specifically designed for the Arabic language to conduct sentiment analysis study on Arabic materials. The pretrained English-Arabic translation model was used to enhance data. For the purpose of creating the new augmented data, the dataset was translated into English and then back to Arabic. A stacking method was used to get the results. The accuracy score of the BERT-based ensemble learning is 96%, according to the model.

Additionally, M. G. Sousa et al. (2019) suggested using bidirectional encoder representations from transformers BERT to analyse news articles' sentiment and offer relevant data for stock market decision-making. Vectorization was carried out using bag-of-words (bow) and term frequency inverse document frequency (tfidf) representation When compared to naive Bayes, support vector machines (SVM), and TextCNN, the dataset exhibited 82.5% accuracy for the manually labelled positive, neutral, and negative stock articles.

The main objective of this project is to provide answers to the following research questions:

1. How can a company's reputation be accurately and thoroughly assessed using sentiment analysis based on customer reviews?
2. Which deep learning or machine learning models do better at appropriately categorizing the sentiments represented in customer reviews?
3. Do the outcomes of the machine learning and deep learning models reflect the Amazon company’s reputation in any way?

METHODOLOGY

Data Collection

The dataset used for this study is Amazon men and women shoes reviews which taken from kaggle.com in a csv file format with a usability of 9.71. After importing the necessary libraries, loading the dataset was the initial step. Python Pandas was used to read the dataset and it shows customer reviews of 856 men and 374 women shoes giving a total of 1230 brands of shoes. It has 1230 multiple review rows which gives a total of 9958 review rows of shoe brands and 8 columns consisting of ‘title, price, rating, total reviews, product description, reviews, reviews rating and shoe type.

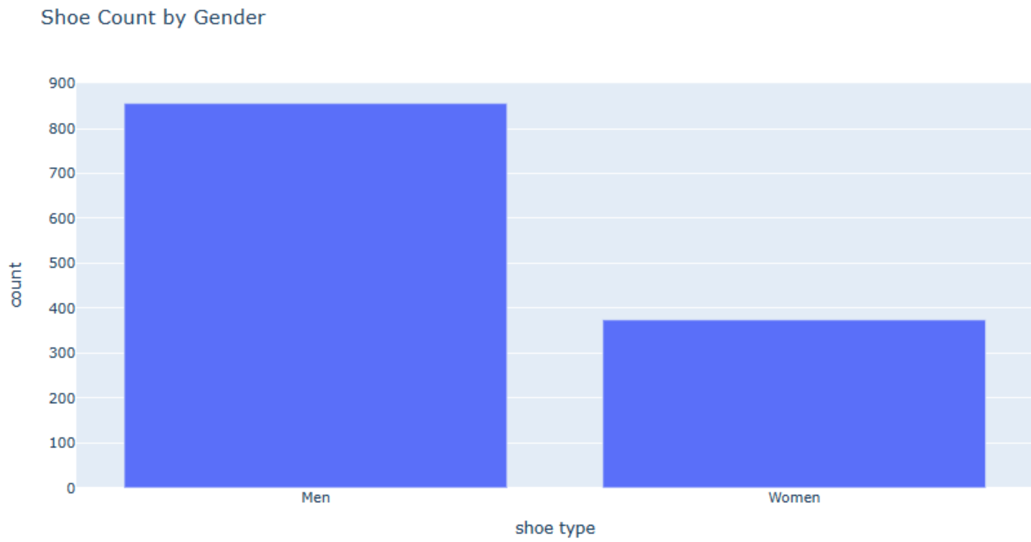

Fig 1. Number of men and women shoe types present in the dataset

Exploratory Data Analysis (EDA)

The dataframe was filtered to the required columns (reviews and reviews ratings) needed for the sentiment analysis and the multiple reviews rows was broken down into 9958 rows to enhance accuracy of the results. The reviews rating was classified into positive and negative sentiments using >=3 as positive and <3 as negative for the purpose of this study. This classification is due to the fact that the percentage of neutral class (3) is 8.67% when compared with positive and negative classes. Therefore, the neutral class (3) was merged to form part of the positive class (Fig.3). This gives a total number of 7660 positive reviews and 2298 negative reviews in the dataset. Exploratory Data Analysis was carried out on the dataset to visualize the dataset in order to have a better understanding of the characteristics of the features of the dataset. This will enable me to check for noisy features and decide on the best approach to manipulate the dataset. This includes graphical representation of the distribution of the reviews ratings and reviews rating counts, unique symbols, word cloud, frequency of words, graph of men and women shoes, average length of words, heat map, percentage of positive and negative reviews ratings.

1

[1] https://www.kaggle.com/datasets/daishinkan002/men-women-shoes-reviews

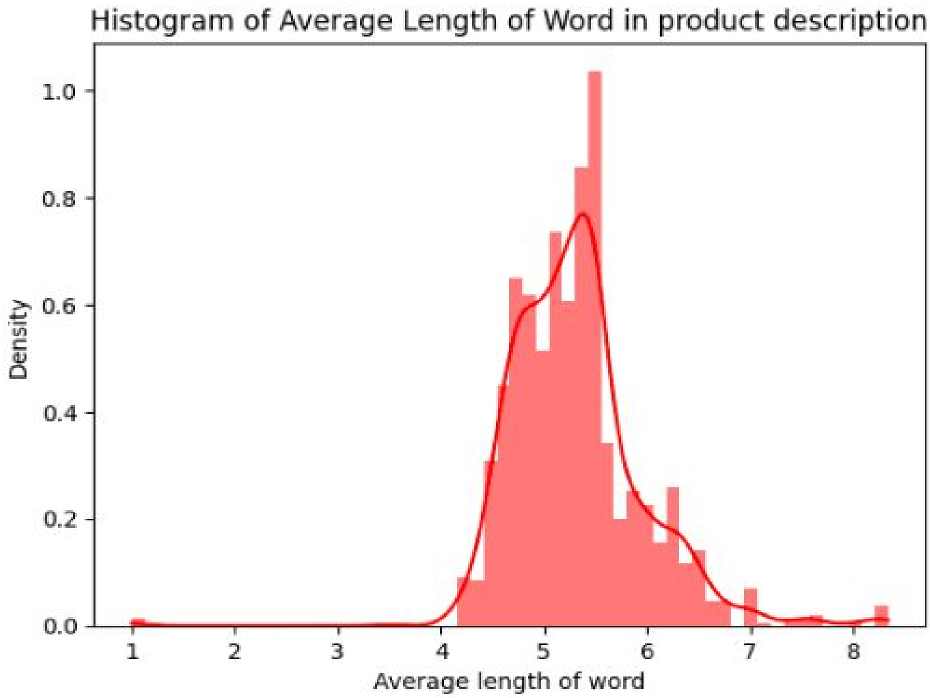

Fig 2. Average length of words in the product description

	REVIEWS	REVIEW RATING	SENTIMENT
0	Not happy with product	1	Negative
1	Its not as expected	1	Negative
2	Average Product	3	Positive
3	Pic more besutiful	3	Positive
4	Got damage product. But quality is average to……	3	Positive
……			
9954	Excellent product	5	Positive
9955	Nice shoes	5	Positive
9956	nice	5	Positive
9957	Asics shoes are the best	5	Positive

Table 1. Sample number of rows in the dataframe with their sentiments after the multiple reviews rows were split

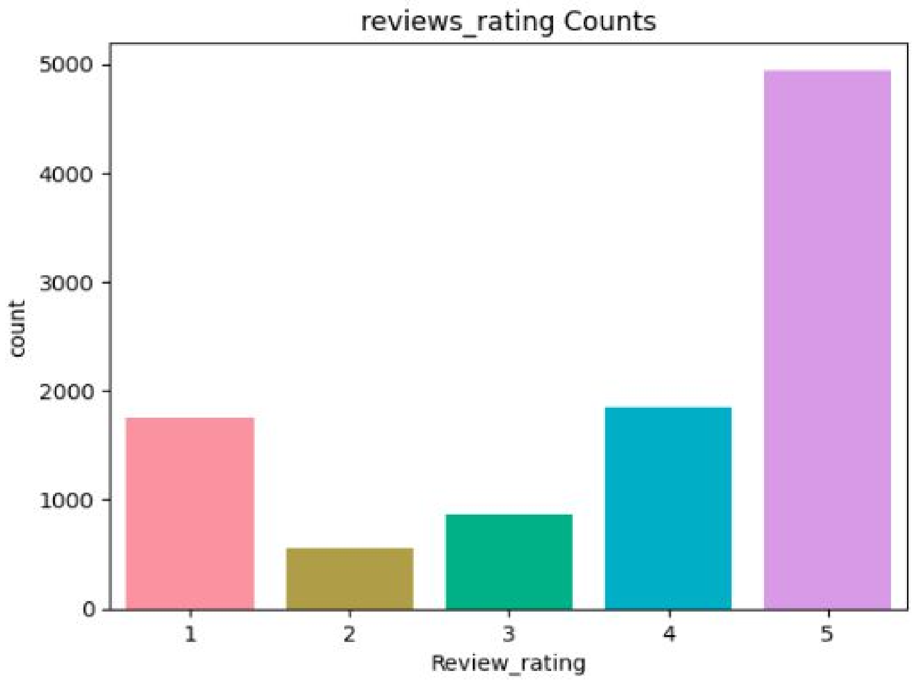

Fig 3. Distrobution of reviews ratings counts in the dataset

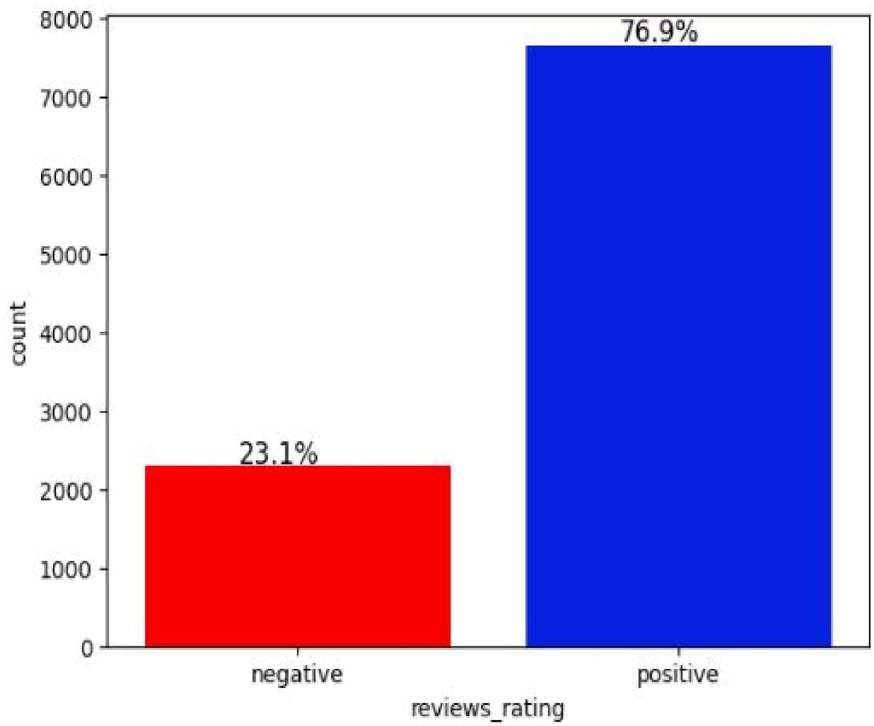

Fig 4. Percentage of the negative and positive reviews discovered in the dataset after classification.

Data Cleaning and Preprocessing

Noisy features were discovered in the dataset during the EDA and data cleaning. These include features with 'verified purchase' and 'report abuse' which randomly formed part of the customer reviews. Features with contractions, emoji, non-English words, stop words, capital letters were also present in the dataset and it became important to clean the dataset before running it on the model to improve accuracy and minimize errors in the algorithms. Contractions was removed from the dataset using python contraction library to expand combination of words (it's – it is). The words were broken down into tokens using word tokenizer. Non-Ascii characters was removed including, emoji, symbols non-English alphabets and words using unicodedata. normalize. The other data cleaning steps used in this study includes removal of punctuations, lowercasing the reviews for uniformity, removal of digits, stop words removal and lemmatization. Lemmatization for word normalization was used over stemming because lemmatization normalizes words without making the words lose their original meaning. Removal of the verified purchase (647) and report abuse (428) reduced the review rows to 8816.

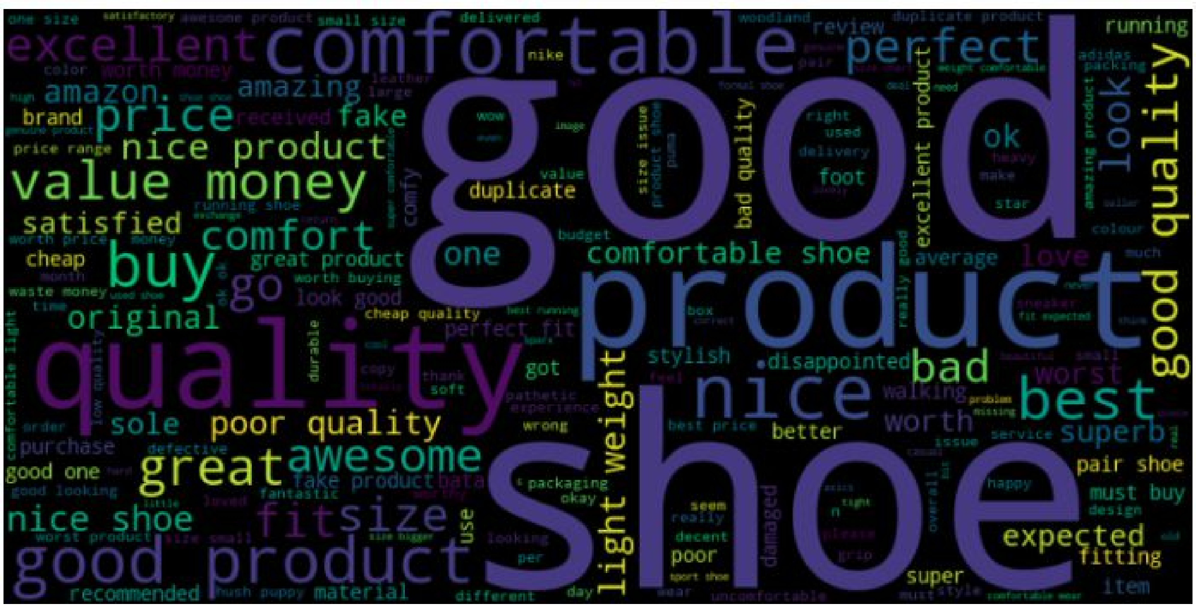

Fig 5. WordCloud of the cleaned customer reviews

Furthermore, the reviews sentiments were mapped into 1 and 0 where positive sentiments were assigned 1 and negative sentiments was assigned 0 as only numeric values can be fed into the models.

Feature Extractions and Word Embedding

For this study the features of the reviews were converted to vectors using word embedding to transform the features of the reviews to numbers in order to allow the machine learning models to carry out the sentiment analysis on the customer reviews. The features for the modelling was extracted using Bag of Words(CountVectorizer) and Term Frequency - Inverse Document Frequency (TF-IDF**).** Bag of Words (BOW) uses the most frequent of words used by customers in the reviews for the word embedding while TF-IDF gives importance to the frequent and non-frequent words used by customers' in the reviews. 500 most frequent words in the customer reviews was used in the Bag of Words methods in order to speed up the processing time.

Balancing the dataset and splitting into test and train datasets

It was discovered that the dataset contains large number of positive reviews compared to the negative reviews so it became important to balance the dataset in order to improve accuracy of the training models and remove bias in the models predictions. This process was carried out using SMOTE from imbalanced learn. The dataset was split into Test and Train datasets with ratio 30% to 70% respectively with 2645 test dataset and 6171 train dataset to be explored in the models.

S/NO	DATASET	COUNTER
1	Original dataset shape	({1: 6658, 0: 2158})
2	Resampled dataset shape	({0: 6658, 1: 6658})

Table 2. Balanced datasets used for the modelling

Modelling

Deep learning and machine learning methods were used for modelling in this study. Firstly, the conventional machine learning methods, i.e., using test datasets for the Random Forest approach and Support Vector Machine (SVM). Random forest uses targeted variables for text classification in machine learning. In a random forest classification, different decision trees are produced using distinct random subsets of the data and characteristics. The Random Forest model was explored on the BOW and TF-IDF. Support Vector Machine (SVM) is used to minimize an error by iteratively generating the best hyperplane by partitioning a dataset into classes to discover a maximum marginal hyperplane (MMH), which is the fundamental notion behind SVM. The SVM model was also examined on the BOW and TF-IDF. These machine learning techniques were selected because they work well with structured data and sentiment analysis tasks.

Later on, the deep learning algorithms, specifically focusing on Distil BERT and Long Short-Term Memory (LSTM) were used to train the datasets. Recurrent neural networks (RNNs) of the LSTM type are excellent at processing sequence data, which makes them a good option for handling textual data. A maximum word length of 16 was observed, and sequences were padded to a fixed length of 7 words for input into the sparse matrix in order to balance the array.

The LSTM architecture was used to train the model then, the application of Distil BERT, a simplified version of the BERT transformer model was also explored. This is because it is lighter and could train 60% quicker than other BERT models. When constructing the Distil BERT model, the Adam optimizer was consistently chosen. Interestingly, a setting of 10 was used for gradient accumulation, indicating a method to build up gradients over several batches prior to executing a weight update. This approach has the potential to be beneficial for optimizing model training.

RESULTS

At the end of the research, the following observations were obtained using the machine learning and deep learning models. The results provide metrics to show the model performance using:

- **Precision**: determines the proportion of the dataset's positive or negative attributes that are genuinely positive or negative based on predictions.
- **Recall**: counts the number of real positive or negative traits that the model accurately predicted.
- **F1 score**: measures the mean of the combination of precision and recall.
- **Support**: shows the number of actual positive or negative occurrences in the dataset.
- **Accuracy**: measures the total number of correct predictions made by the model.

Machine Learning Models

1. **Random forest**

	precision	recall	f1-score	support
0	0.77	0.63	0.69	647
1	0.89	0.94	0.91	1998
accuracy			0.86	2645
macro avg	0.83	0.78	0.80	2645
weighted avg	0.86	0.86	0.86	2645

Fig 6. Result of Random Forest model on Bag of Words(BOW)

The result of Ransom forest model application on BOW shows the precision values of 77% for class 0(negative) and 89% for class 1(positive), the classification model performs well and shows excellent accuracy in properly identifying instances of each class. Class 0 recall ratings of 63% and class 1 recall values of 94% indicate that a considerable percentage of real positive cases are captured by the model. Class 0 and class 1 balanced F1-scores of 0.69 and 0.91 respectively highlight how well the model balances recall and precision. The model's ability to make accurate predictions over the whole dataset is further supported by its overall accuracy of 86%, macro and weighted averages around 0.86 shows that performance across classes is well-balanced. (Fig.6).

	precision	recall	f1-score	support
0	0.83	0.62	0.71	664
1	0.88	0.96	0.92	1981
accuracy			0.87	2645
macro avg	0.86	0.79	0.81	2645
weighted avg	0.87	0.87	0.87	2645

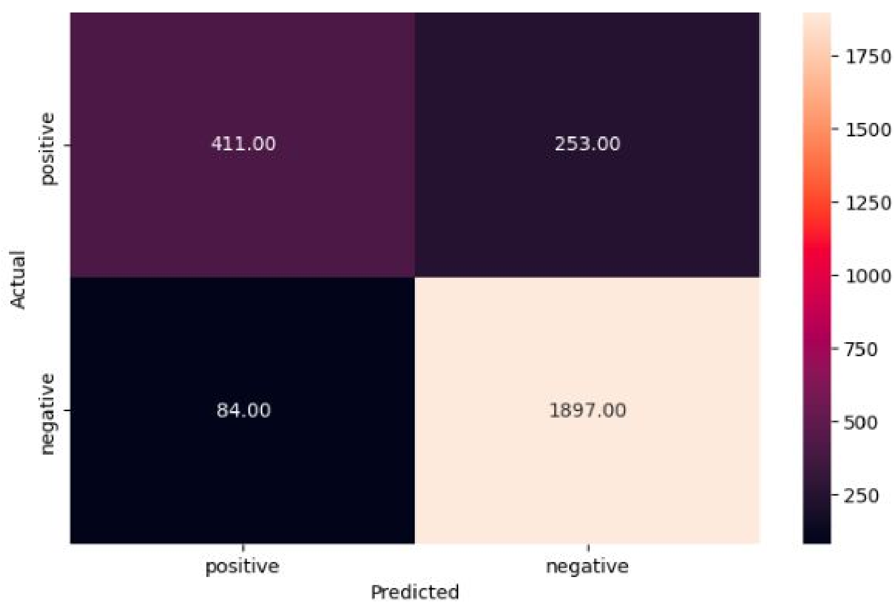

Fig 7. Result of Random Forest model on TF-IDF

With precision values of 94% for both class 0(negative) and class 1(positive), the Random forest model performs well and shows excellent accuracy in Class 0 recall ratings of 91% and class 1 recall values of 96% indicate that a considerable percentage of real positive cases are captured by the model on TF-IDF (Fig 7). Class 0 and class 1 balanced F1-scores of 0.93 and 0.95 respectively highlight how well the model balances recall and precision. The model's ability to make accurate predictions over the whole dataset is further supported by its overall accuracy of 94%, and macro and weighted averages around 0.94 show that performance across classes is well-balanced. All things considered, these findings imply that the classification model is trustworthy and accurate in differentiating between the two classes

2. **Support Vector Machine(SVM)**

   The SVM model evaluation metrics for TF-IDF (Fig 8) show that the model maintains a precision and recall value in negative class 0 predicts 83%, 52% of the time, and captures 64% of the real class 0 cases. In a similar vein, positive class 1's precision, recall, and F1-score are 86%, 96%, and 91%, respectively. The model has 85% overall accuracy rate in both classes, demonstrating its high predictive power. The weighted and macro averages, which hover around 85%, highlight the dataset's overall performance being balanced and steady**.**

	precision	recall	f1-score	support
0	0.82	0.52	0.64	664
1	0.86	0.96	0.91	1981
accuracy			0.85	2645
macro avg	0.84	0.74	0.77	2645
weighted avg	0.85	0.85	0.84	2645

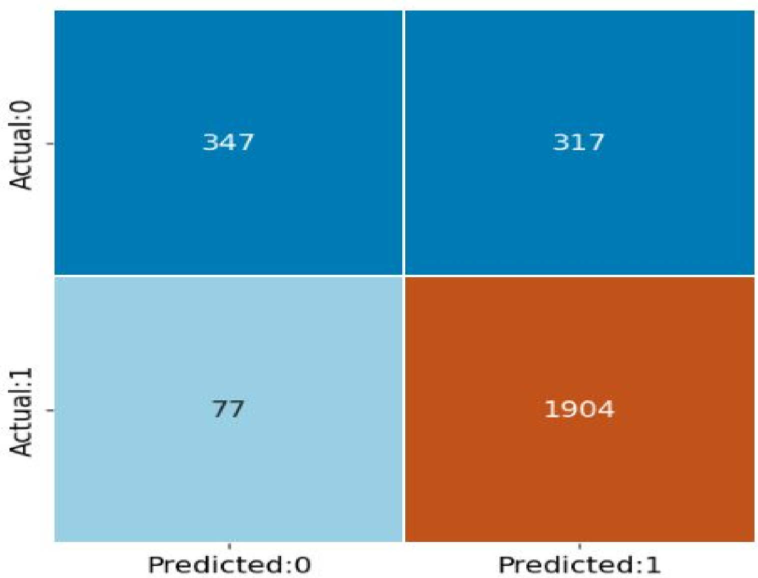

Fig 8. Result of the SVM model on TF-IDF

The SVM model evaluation metrics for BOW (Fig 9) show that the model maintains a balanced precision-recall trade-off, is accurate in negative class 0 predicts 83% and 55% of the time, and captures 68% of the real class 0 cases. Similarly, positive class 1's precision, recall, and F1-score are 87%, 96%, and 91%, respectively. The model has 86% overall accuracy rate in both classes, demonstrating its high predictive power.

```
              precision    recall  f1-score   support

           0       0.83      0.55      0.66       647
           1       0.87      0.96      0.91      1998

    accuracy                           0.86      2645
   macro avg       0.85      0.76      0.79      2645
weighted avg       0.86      0.86      0.85      2645
```

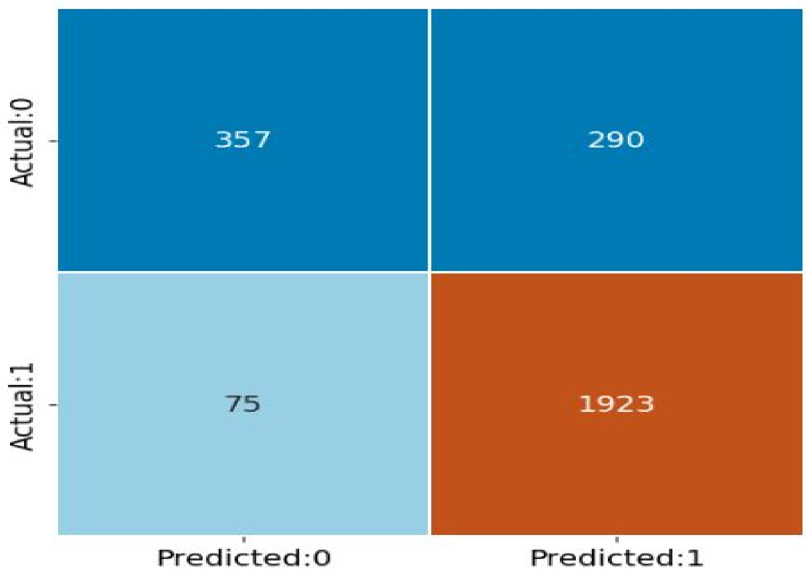

Fig 9. Result of the SVM model on BOW

MACHINE LEARNING MODEL EVALUATION RESULTS

S/NO	MODEL		ACCURACY(%)
1	Random Forest	BOW	86
		TF-IDF	87
2	SVM	BOW	86
		TF-IDF	85

Table 3. Evaluation results of the Machine learning models

Deep Learning Models.

1. Long Short Term Memory Model(LSTM)

   The training accuracy of the LSTM model rises from 79.98% to 90.50% while the training loss decreases from 45.19% to 24.21%. Validation results show accuracy from 85.43% to 85.345 while validation loss indicates 34.66% to 39.19%. This suggests that the model is getting better at predicting the dataset's positive and negative sentiments. The LSTM model has a little positive model bias, which indicates that it performs marginally better on the train dataset than the validation dataset. Given that the model accuracy for the train dataset is marginally greater than the model accuracy for the validation dataset, this shows that the model might be overfitting to the training data**.**

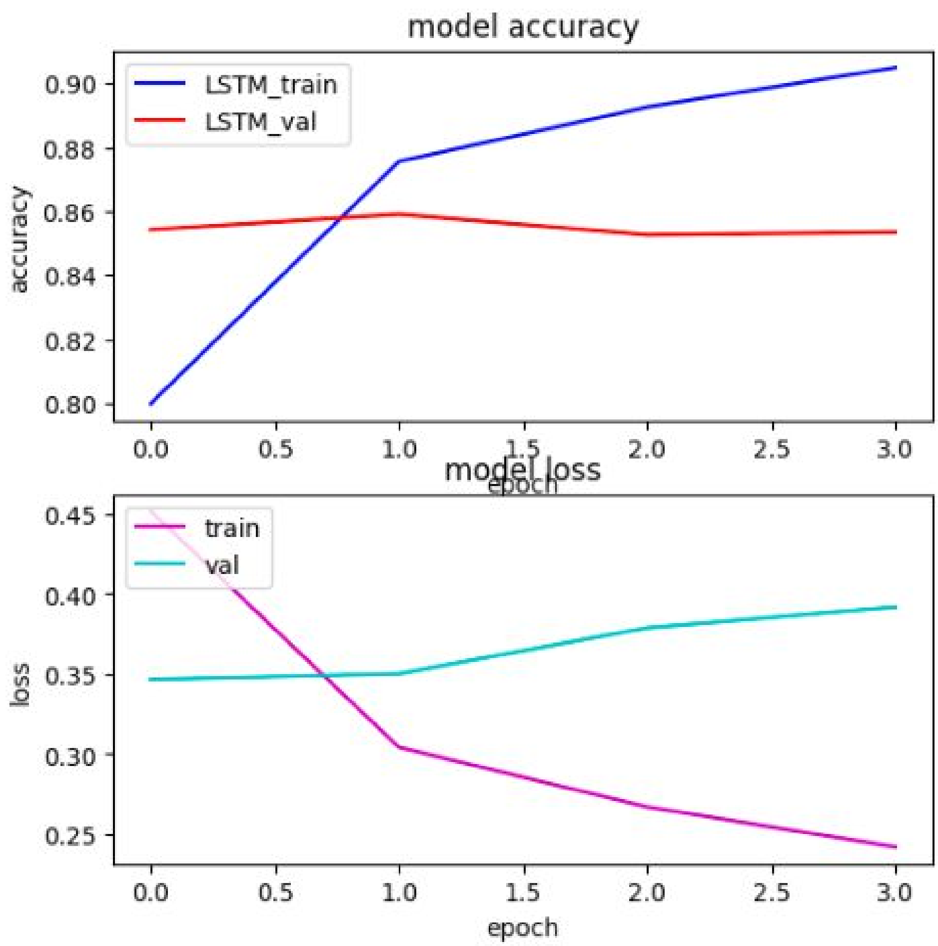

Fig 10. The graphs showing the result of LSTM Model on the training dataset

2. DistilBERT

The sentiment analysis DistilBERT model result demonstrates the model's excellent performance. After 220 training steps, validation accuracy of 80% indicates that the model can accurately categorize both positive and negative sentiments with a high degree of accuracy and the model is still learning with no sign of overfitting**.** (Table 4). The result also demonstrate that the model is capable of performing text classification tasks like customers' reviews.

DEEP LEARNING MODEL EVALUATION RESULTS

S/NO	MODEL	VALIDATION ACCURACY(%)
1	LSTM	85
2	DistilBert	80

Table 4. Evaluation results of the Deep learning models

DISCUSSION

The aim of this study is to examine how Sentiment Analysis can be used to provide accurate and comprehensive assessment of a company's reputation using customers' reviews. By responding to the research questions in this study, the outcomes from the models utilized in this study will be examined.

Q1: How can a company's reputation be accurately and thoroughly assessed using sentiment analysis based on customer reviews?

The results generated from this study as shown in table 3 and table 4, Amazon's shoes reviews can be accurately and thoroughly assessed by analyzing the sentiments expressed in customers' reviews. This analytical method can therefore be generalized to draw conclusions on the assessment of customers' perception about a company's reputation. The high precision and recall values indicates that the models are able to make accurate predictions across the negative and positive sentiments. F1 score also shows indication of a good balance between the precision and recall. The overall accuracy result of the models shows that there is no bias towards positive sentiments. However, the overfitting the LSTM model in fig. 9 shows that there is a slight positive bias towards the training datasets.

Consumers of today rely their decisions about what to buy on information they may obtain on online retail sites. Their purchasing behaviour may be significantly impacted by the information posted on internet sites. Because they are unable to engage in person with products or sellers, online shoppers often rely on platform information, such as product quality and company reputation, to gain an indirect understanding of a product before making a purchase. (Da Yeon Kim, Sang Yong Kim, 2022).

Q2: Which deep learning or machine learning models do better at appropriately categorizing the sentiments represented in customer reviews?

With good precision, recall, and F1-scores for both positive and negative sentiments, the Random Forest model performed excellently showing accuracy of 86% and 87% for BOW and TF-IDF on the test datasets. The Support Vector Machine (SVM) model exhibits 86% and 85% in BOW and TF-IDF as shown in table 3. In table 4, the LSTM models exhibit validation accuracy 85%. The validation accuracy of the DistilBERT model is a high 80%. Overall, from the model results, it appears that both deep learning and machine learning models do a good job of classifying sentiments based on this study. The results show that Random Forest machine learning model and LSTM deep learning model are excellent in categorizing sentiments expressed in the customer reviews. However, the choice between these models may be influenced by factors including the task's particular needs, processing capacity, optimization, tuning hyper-parameters as well as dataset size can influence the model performance.

Q3. Do the outcomes of the machine learning and deep learning models reflect the Amazon company’s reputation in any way?

From table 3 and 4, both deep learning and machine learning models perform well when it comes to classifying the sentiments from customer reviews. Both the deep learning models, such LSTM and DistilBERT, and the machine learning models, like Random Forest and SVM, exhibit great accuracy and well-balanced performance. In this study, the LSTM model confirms the proficiency giving validation accuracy of 85% when compared with previous work by Jihua Cao et al. (2023) where the LSTM model generated a validation accuracy of 89.52% This implies that LSTM is proficient in managing tasks related to sentiment analysis which overall determines the company’s reputation. The outcome of this study reflects that Amazon has a good reputation in the shoes line of the company’s business.

CONCLUSION

The sentiment analysis outcomes from this study's machine learning and deep learning models provides insightful information about how well various methods categorize sentiments. With balanced F1-scores for both positive and negative classes, the Random Forest models demonstrate remarkable precision (0.89, 0.88) and recall (0.94, 0.96) rates for positive reviews when applied to both the Bag of Words (BOW) and TF-IDF representations (fig 6 and fig 7). The precision-recall trade-off exhibited by the Support Vector Machine (SVM) model is balanced, and it achieves good accuracy rates in predicting sentiments, both positive and negative. Weighted and macro averages reflect the overall well-balanced performance, which highlights the model's accuracy in categorizing sentiments.

The Deep Learning models show excellent performance, especially the Long Short-Term Memory (LSTM) models with an 85% validation accuracy. The model shows a minor bias towards the training dataset. These results imply that the LSTM models are skilled at identifying sentiment trends and producing precise forecasts. With a validation accuracy of 80%, the DistilBERT model, an advanced transformer-based architecture, has exceptional performance. This demonstrates the model's capacity to classify sentiments accurately and highlights the value of pre-trained language models in sentiment analysis applications.

Understanding the emotions expressed in the reviews provides important information for evaluating and enhancing the company's products or services. It also provides insights into the preferences, opinions, and general level of customer satisfaction. Positive or negative product reviews can influence sales, which determines a company's success or failure in the marketplace. This study confirms that sentiment analysis is excellent in analyzing a company’s reputation. As a crucial business intelligence tool, sentiment analysis forecasts a company's overall reputation and offers unbiased customer insights, big customer data analysis, and real-time outcomes that help businesses enhance their products and services.

Future work could be to extend this study across the entire Amazon corporation considering different languages and look into the benefits and problems associated with cross-lingual sentiment analysis. Applications in international or multilingual situations may find this useful. The use of transfer learning by utilizing pre-trained word embeddings like Word2Vec or GloVe. Optimizing this embeddings for the sentiment analysis task may help the model perform better, particularly when it comes to identifying subtle linguistic patterns.

REFEENCES

Chouikhi, Hasna & Jarray, Fethi. (2023). BERT-Based Ensemble Learning Approach for Sentiment Analysis. 10.1007/978-3-031-35924-8_7

Davoodi, Laleh & Mezei, József. (2022). A Comparative Study of Machine Learning Models for Sentiment Analysis: Customer Reviews of E-Commerce Platforms. 217-231. 10.18690/um.fov.4.2022.13.

Da Yeon Kim, Sang Yong Kim. (2022). The impact of customer-generated evaluation information on sales in online platform-based markets, Journal of Retailing and Consumer Services, Volume 68, 2022,

Falasari, Anisa & Muslim, Much. (2022). Optimize Naïve Bayes Classifier Using Chi Square and Term Frequency Inverse Document Frequency for Amazon Review Sentiment Analysis. Journal of Soft Computing Exploration. 3. 31-36. 10.52465/joscex.v3i1.68.

Huggingface. DistilBert. https://huggingface.co/docs/transformers/model_doc/distilbert

Jagdale, Rajkumar & Shirsath, Vishal & Deshmukh, Sachin. (2019). Sentiment Analysis on Product Reviews Using Machine Learning Techniques: Proceeding of CISC 2017. 10.1007/978-981-13-0617-4_61.

JihuaCao, JieLi, MiaoYin, andYunfengWang.2023. Online Reviews Sentiment Analysis and Product Feature Improvement with Deep Learning. ACMTrans.AsianLow-Resour.Lang.Inf.Process.22,8,Article203

Kovalova, Hanna & Ali, Adil & Zamlynskyi, Victor. (2021). Business reputation of the company as one of the most important components of the company's success. Economics. Finances. Law. 29-32. 10.37634/efp.2021.6(2).6.

M. G. Sousa, K. Sakiyama, L. d. S. Rodrigues, P. H. Moraes, E. R. Fernandes and E. T. Matsubara, "BERT for Stock Market Sentiment Analysis," *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, Portland, OR, USA, 2019, pp. 1597-1601, doi: 10.1109/ICTAI.2019.00231.

Mutinda, James. (2022). Sentiment Analysis on Text Reviews Using Lexicon Selected-Bert Embedding (LeBERT) Model with Convolutional Neural Network. 10.21203/rs.3.rs-2330887/v1

P. Manjula, Neeraj Kumar and A (2021). Customer Sentiment Analysis Using Cloud App and Machine Learning Model10.1007/978-981-15-7990-5_32

Random Forest Classification with Scikit-Learn. https://www.datacamp.com/tutorial/random-forests-classifier-python

Seoyeon Hong & Matthew Pittman (2020) eWOM anatomy of online product reviews: interaction effects of review number, valence, and star ratings on perceived credibility, International Journal of Advertising, 39:7, 892-920, DOI: 10.1080/02650487.2019.1703386

Shantanu Tripathi (2021). Men_ Women_ Shoes_Reviews https://www.kaggle.com/datasets/daishinkan002/men-women-shoes-reviews

Sharma, S. K., Chakraborti, S. and Jha, T. 2019. Analysis of Book Sales Prediction at Amazon Marketplace in India: A Machine Learning Approach. Information Systems and eBusiness Management, 17(2–4), pp. 261–284

Support vector machines with Scikit-learn. https://www.datacamp.com/tutorial/svm-classification-scikit-learn-python