Enhancing Recommender Systems with Sentiment Analysis based on E-Commerce.

Naydel Muscat

Supervisor: Jane Cilia Debono

June - 2024

A dissertation submitted to the Institute of Information and Communication Technology in partial fulfilment of the requirements for the degree of BSc (Hons) Business Analytics

Authorship Statement

This dissertation is based on the results of research carried out by myself,it is my own composition, and has not been previously presented for any other certified or uncertified qualification.

The research was carried out under the supervision of Jane Cilia Debono

. . . .

Date   Signature

Copyright Statement

....	....
Date	Signature

Acknowledgements

I would like to thank my Supervisor, Ms Jane Cilia Debono, for guiding me on the right path, providing me with assistance and support throughout the research I seeked to accomplish.

I would like to also thank my family, friends and my partner for supporting me throughout these stressful months and encourage me.

Abstract

This paper investigates the use of machine learning and deep learning techniques in sentiment analysis. Hence, applying sentiment analysis in e-commerce recommender systems. Such techniques used to achieve this include: Logistic Regression, Support Vector Machine (SVM), Random Forest, Convolutional Neural Networks (CNN); and Long Short-Term Memory (LSTM). The Amazon Product dataset specifically the Cell Phones and Accessories 5-core, was utilised for this research. This dataset comprises of reviews that users gave to a product, spanning over the years from 1996 through 2014. The primary focus of this study is on evaluating how sentiment analysis can improve the accuracy and prediction of recommendations to users. Three types of recommender systems were developed all utilising an SVD model, a collaborative filtering approach. The baseline system uses only the ratings feature. The sentiment analysis model makes use of sentiment scores. Whereas, the combined system integrates both ratings and sentiment scores. Metric results were necessary to evaluate the performance of the models. For sentiment analysis: Accuracy, Precision, F1-score, Recall, and AUC score were utilised. RMSE and MAE were employed for the recommender system. The SVM model outperformed all the other models. Thus, achieving an accuracy of 0.93, a precision of 0.93, a recall of 0.93, an F1-score of 0.92; and an AUC score of 0.95. The sentiment-based recommender system achieved an RMSE of 0.43 and an MAE of 0.42 outperforming the baseline system, which had an RMSE of 1.20 and an MAE of 0.87. The combined model further improved the RMSE to 1.14 and the MAE to 0.79.

Keywords: Sentiment Analysis, Recommender System, Machine Learning, Deep Learning, E-Commerce.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ML	Machine Learning
RS	Recommender System
SA	Sentiment Analysis
NB	Naïve Bayes
SVM	Support Vector Machine
LR	Logistic Regression
SVD	Singular Value Decomposition
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
RMSE	Root Mean Square Error
MAE	Mean Absolute Error

Chapter 1: Introduction

1.1 Research Background

E-commerce has revolutionised the way consumers interact with businesses, offering a vast range of products and services. However, the overwhelming amount of information available online poses challenges to users with personalised and relevant recommendations. Therefore, the use of recommender systems have emerged to solve this problem. Thereby, enhancing user experience by suggesting items that align with their individual preferences and needs. Traditional recommender systems often rely on numerical ratings. Yet, this approach often falls short on capturing successful recommendation based on the user preferences. This leads to the necessity of exploring new approaches to improve the accuracy and user satisfaction of these systems.

Sentiment Analysis is one of the promising solutions to address this limitation of the e-commerce recommender systems. It involves the extraction and analysing of textual data, such as the user reviews. Hence, incorporating this method into these systems can assist in achieving a deeper understanding in user preferences, and will lead to more accurate recommendations. In addition, applying sentiment analysis not only enhances traditional recommendation techniques but also improves the user experience as it considers the user feedback.

1.2 Aims and Objectives

This project's aim is to determine and explore how machine learning sentiment analysis can be implemented into recommender systems based on e-commerce. Additionally, the main objective of this research is to investigate the traditional recommender systems when applying the sentiment analysis feature. Hence, the final implementation generates high-quality recommendations for a specific user based on their preferences.

1.3 Hypothesis and Research Questions

The hypothesis of this study is to show that applying sentiment analysis into traditional recommender systems as a feature; it will improve the overall performance of the system to recommend to users based on their preferences. To test the hypothesis, the following research questions are addressed:

1. Which model performs better in terms of accuracy, precision, recall, F1-score, and AUC score?
2. Are deep learning algorithms better than traditional machine learning algorithms in terms of sentiment analysis?
3. Can sentiment analysis boost the performance of traditional recommender system?

1.4 Research Outline

The first chapter introduces the research topic. It outlines the significance of enhancing recommender systems with sentiment analysis in the context of e-commerce. This sets the stage for the study by discussing the limitations of traditional recommender systems and potential enhancements when applying sentiment analysis into recommender systems.

The Literature Review provides a summary of existing research on e-commerce recommender systems and sentiment analysis. This chapter contains: the types of recommender systems, types of sentiment analysis, and sentiment analysis in recommender systems. Additionally, identifies gaps in the current study and highlights the need for further investigation.

The third chapter details the employed research methodology for this project. It highlights: the research problem, solution, research process, data collection and tools, prototype development and ethical issues.

The following chapter then discusses the analysis and findings of the study. Furthermore, it compares the outcomes with existing literature; addresses the research questions and hypothesis.

The last chapter summarises: the research findings and results, future recommendations, limitations; and identifies areas for further research.

1.5 Conclusion

The next chapter, Literature Review will give a solid foundation for this current research. It will cover a variety of topics including e-commerce, recommender systems, and sentiment analysis.

Chapter 2: Literature Review

This chapter will present a literature review pertaining to the selected research topic. It will cover the topics of e-commerce, recommender systems, and sentiment analysis.

2.1 E-Commerce

Electronic commerce, also known as e-commerce, is an online business that has a wide range of activities for products and services to be sold. In the e-commerce sector payments occur electronically rather than having people interact with each other by direct physical contact. Hence, this type of business is usually associated with buying, selling goods and services over the internet. In addition, any transactions involving ownership or rights in utilising the products or services through electronic devices communicate interactively over network channels. The definition of e-commerce is not so broad. This is due to the fact of the recent developments in this new and revolutionary business. A simple explanation can be, the use of electronic communications and digital information. Processing technology in business transactions create, transform, and redefine relationships for value creation. This applies both, to organisations and between organisations and individuals. The different parts of e-commerce include:

1. Business-to-business (B2B)
2. Business-to-consumer (B2C)
3. Business-to-government (B2G)

4. Consumer-to-consumer (C2C)
5. Mobile commerce (m-commerce) [1]

2.2 Recommender Systems

E-commerce plays a crucial role in today's business environment by giving customers a vast amount of information, which improves the user experience. Hence, to overcome the overwhelming amount of information, one solution is the use of recommender systems. Recommender systems provide personalised recommendations to users and better customer satisfaction. The use of these systems has increased due to their popularity with their good performances in achieving satisfaction. Ongoing research enhances this technology's performance and accuracy in providing recommendations. As a result, various fields such as entertainment and education have been adapting the use of recommender systems [2]. Such techniques of recommender systems include: content-based filtering, collaborative filtering (User-Item, User-User), context-based and hybrid filtering. Furthermore, deep learning is emerging and increasingly integrated into these systems [3].

2.2.1 Recommender Systems in E-Commerce

Recommender systems have rapidly gained popularity in the e-commerce industry. This is due to the fact of their remarkable capacity of enhancing personalisation and automated customisation. An increase in sales on e-commerce websites has been observed as businesses adapt to visitors' needs by integrating their input into their websites; gradually converting them into potential customers over time.

In addition, recommender systems can assist websites to advertise extra products, by bundling closely related things together and increasing customer loyalty.

Hence, this technique enhances e-commerce sales in three particular ways. Customer loyalty can be accomplished by demonstrating to customers that the business takes time to understand their needs and to find more information to meet those demands. As seen with the products, website structure and presentation are all tailored to the requirements and preferences of the user. A customer is more likely to return to these websites than a competitor's this is because they are accustomed to it and do not have to go through the same learning process to adjust their recommendations. Customers typically remain with a website they are familiar with, even if the competition is providing a better or comparable experience [4]. Browsers into buyers, visitors often look over a website without ever making a purchase. Therefore, with the use of recommender systems users can locate products they wish to purchase. Cross-sell, recommender systems enhance cross-selling by recommending more items for the client to buy. For instance, based on the things that are currently in the shopping cart, a website may suggest more products throughout the checkout process [5].

2.2.2 Collaborative Filtering

Collaborative filtering is one of the most popular recommender systems successfully used in the field of e-commerce. This technique is a domain-independent prediction method. It focuses solely on identifying the user's closest neighbours who either purchased the same products or gave similar ratings to the same products as the target user [6]. Moreover, it is employed for content like music and movies that are difficult to categorise using metadata. This approach makes use of a user-item matrix, a database designed specifically to record the user's preferences for various products [7]. Therefore, computations are needed to detect similarities between a user and the target user in order to determine the user's nearest neighbour.

One of the most used algorithms for this method is cosine similarity. Cosine similarity suggests products bought by the neighbour after locating the target user's closest neighbour, even if the target user hasn't made any purchases [6].

This technique is then divided into two models: memory-based and model-based. A memory-based collaborative filtering model consists of having items that were rated by the user. Then, these are used to find neighbours that share the same interests.

Memory-based is then split into two more techniques user-based and item-based. Model-based technique considers the previous ratings to train the model in order to achieve a better performance for the collaborative filtering approach. Algorithms that have been used in model-based include Association rule, Clustering, Decision tree, Artificial Neural Network, and regression [7].

P. H. Aditya, I. Budi, and Q. Munajat investigated a comparison of model-based and memory-based collaborative filtering algorithms in their study. The techniques that were implemented used PT X, a well-known company in Indonesia. This dataset comprised of: 50,000 Users, 95,468 Products; and 290,060 Transactions. In addition, it was then separated into three datasets and pre-processed to exclude null and redundant values. Since the dataset lacked rating information, a Naïve Bayes approach was applied for model-based and a nearest neighbour algorithm was employed for the memory-based. Data was split 70% for training and 30% for testing, in which the input of the trained data was utilised to generate both model-based and memory-based recommendations.

The testing set was implemented to evaluate the recommender system, where the F1 metric evaluator was employed for offline testing. Moreover, the researchers built an online testing platform that participants may access. In terms of accuracy, the model-based approach outperformed the memory-based approach. Furthermore, the model-based approach received a score of 17.07% in user-based testing, whereas the memory-based technique received a score of 17.20%. This testing showed which model provided users with the most appropriate suggestions [8].

Another study conducted by L. Jiang, et al. a trust-based collaborative filtering algorithm for an E-commerce recommender system. The researchers utilised the Amazon ratings dataset that is presented in table 2.1, which was applied for offline experiments to compare prediction accuracy of various algorithms.

Table 2.1: *Amazon Ratings Dataset*

Number of reviews	34,686,770
Number of users	6,643,669
Number of products	2,441,053
Users with >50 reviews	56,772
Median no. of words per review	82
Timespan	Jun 1995 – Mar 2013

The data was split into 4:1 training and testing sets. The method the researchers used was the slope one algorithm. It was based on the fusion of trusted data and user similarity which has improved the user similarity in recording the rating matrix of the users-items and computing similarity. The newly improved algorithm outperformed the traditional slope algorithm in accuracy.

Table 2.2 shows the MAE and RMSE of the traditional algorithm and the Improved algorithm where R is the trusted ratio of user ratings. As a result, the enhanced algorithm outperformed the standard approach in its results. One main issue that the researchers discovered is the cold-start problem. However, content information assisted in closing the gap between existing items to new items by identifying similarities among them [9].

Table 2.2: *Traditional Slope Algorithm vs Improved Slope Algorithm*

R	>0	>0.5	>0.8	=1	Null
MAE Improved Algorithm	1.189	1.064	0.758	0.598	0.798
MAE traditional Algorithm	1.389	1.214	0.776	0.659	0.967
RMSE traditional Algorithm	0.143	0.146	0.144	0.158	0.049
RMSE Improved Algorithm	0.164	0.164	0.145	0.187	0.057

2.2.3 Content-based Filtering

A content-based filtering (CBF) approach, also referred to as cognitive filtering, recommends items according to the user's profile and derives attributes from the details of previous items that have been reviewed by the user [10]. Several methods are used in CBF to identify similarities between items and generate effective recommendations. It employs Vector Space Model that includes Term Frequency Inverse Document Frequency (TF/IDF) or Probabilistic Models including Decision Trees, Naive Bayes, as well as Neural Networks for modelling the correlation between different items within the corpus. CBF addresses challenges found in collaborative filtering by recommending new items to users even though the user has not rated anything. However, the issues of content-based filtering are that it requires a thorough understanding and description of the characteristics of the items. In addition, for the model to work properly, a well-managed user profile is necessary; this problem is commonly known as limited content analysis [7].

An architecture for a content-based recommender system that focused on related e-commerce products was presented by Nurcahya and Supriyanto. The data came from an Indonesian e-commerce website comprising 868 datasets including food items, animal goods, and agricultural products. The information has been labelled to identify particular items, including vegetables, based on their shape, components, plant species. Cosine similarity and the TF-IDF algorithm have been employed for the proposed model, with recall and precision metrics introduced for system evaluation. As a result, test schemes were developed. The first scheme recommended a product based on the document's name and description, yielding an average recall value of 0.30 and a precision value of 0.59. In contrast, the second scheme had an average recall of 0.84 and a precision of 0.78 for product recommendations based on document name. They discovered limitations as the recommender system was not as successful as it took them ten minutes to generate a recommendation. Also, recommendations were not intended for REAL time [11].

Another paper was proposed by Sabitha, Vaishnavi, Karthik, and Bhavadharini in 2022. A novel User Interaction Based recommender system directed for handling multiple user sources. 120 customer samples were selected from the dataset from the e-commerce website Askboutique so that the model could process them.

The customer's age, gender, purchasing history, and product reviews are included in the dataset. Additionally, the preferences of users regarding the products of a certain website were ascertained through the use of a Content-based filtering model. Moreover, the RBF Neural Network was employed to analyse the inputs and give lists of recommended products based on their computed weights, and the Dempster Shafer (D-S) theory was applied for data integration. The researchers made use of the Precision and Recall rate for the testing of the system. The model resulted in having higher precision rate and greater Recall rate than previous works in generating similar product recommendations to users based on their interactions and interests during online purchases [12].

2.2.4 Sentiment Analysis

Sentiment Analysis (SA) encapsulates a wide range of studies The focal point of these studies is the sentiment analysis technique that is utilised to classify and detect opinions expressed in text. Sentiments can either be: positive, neutral, or negative overall. Several application industries including: social media, corporations, education; and government have discovered considerable benefits when using SA [13]. In addition, such common techniques in sentiment analysis are lexicon-based, machine learning and hybrid, with researchers continually seeking to find more efficient methods to perform tasks at a lower computational cost with higher accuracy. This can be seen in Figure 2.1 which provides an overview of the various methods used in Sentiment Analysis [14].

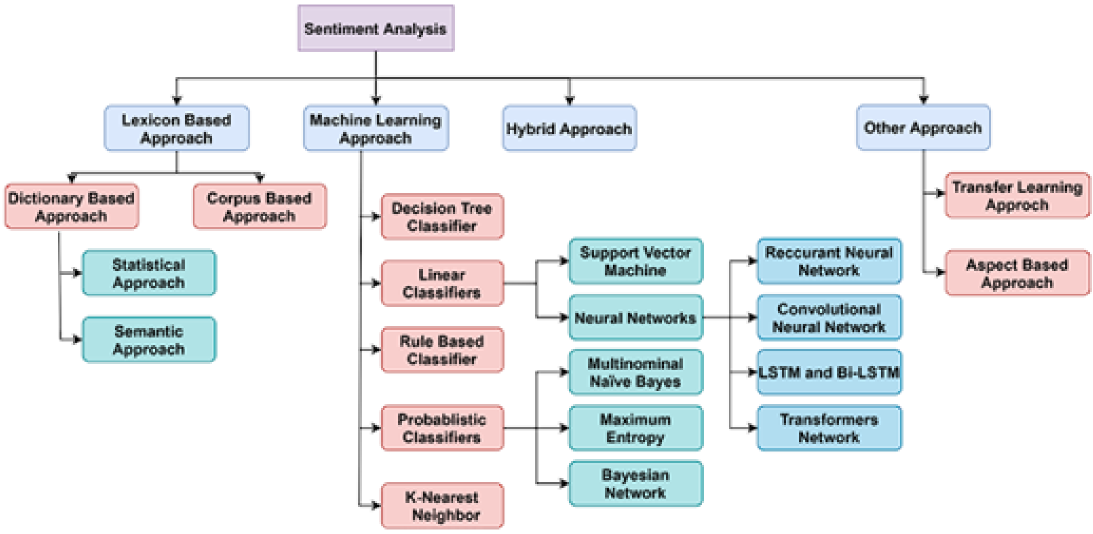

Figure 2.1: *Techniques of Sentiment Analysis*

Lexicon based approach

A lexicon-based approach uses a set of tokens (words), each with a fixed score indicating whether the text is neutral, positive, or negative. This technique is especially effective for sentiment analysis at the sentence and feature levels since it does not require any training data. Hence, making it an unsupervised method. However, one major disadvantage when using this algorithm is that words can have several meanings and senses. Therefore a positive phrase in one domain can be a negative phrase in another [14]. Additionally, lexicon-based approach is divided into two primary methodologies: the Corpus-based and dictionary-based [15] [16].

Machine Learning Approach

Another approach found in sentiment analysis studies is utilising machine learning (ML). ML is an artificial intelligence method that is used for teaching computers using: supervised, semi-supervised, unsupervised; and hybrid approach. Furthermore, the use of machine learning algorithms has emerged in SA as the techniques can capture textual features without requirements of high-level feature extraction. As a result, achieving sentiments on customer reviews when utilising these techniques. Such algorithms of machine learning that have been incorporated in sentiment analysis include:

- **Naïve Bayes (NB)**: The technique Naïve Bayes is used both for classification and training. NB is a Bayesian classification approach based on the Bayes theorem that predicts the probability of a given collection of characteristics being part of any label.

- **Support Vector Machine (SVM)**: The SVM algorithm analysis data using hyperplanes and defines decision limits. Support vector machines are a sort of non-probabilistic supervised learning technology used in classification tasks. The basic goal of an SVM model is to find the hyperplane with the highest possible margin. In addition, several research have shown SVM's in sentiment analysis [14]. Borg and Boldt conducted a project in 2020 on using SVM with the combination of VADER to assess consumer sentiment in emails. This strategy performed well in terms of F1-score and AUC score to predict sentiment patterns in email discussions [17].

- **Logistic Regression (LR)**: This machine learning algorithm evaluates the value of input features in discriminating between positive and negative classes by multiplying each feature by its corresponding weight. This classification approach is intended to discover the key components for this differentiation.

Hence, LR makes use of a probabilistic regression analysis to categorise data [14]

- **Random Forest**: This supervised learning approach is based on the decision tree algorithm. This method aggregates the predictions of several estimators, each constructed using the decision tree. Consequently, when using the multiple estimators increase the strength and boost the model beyond of what a solo estimator would have achieved. This technique is also known as forest. It consists of creating a significant amount of classification trees followed by determining the final decision based on the majority vote among the trees. Overall, having more trees in the forest improves the model's accuracy.

  Hence, this technique has been employed to address text classification challenges, including: hate speech detection, authorship profiling, and sentiment analysis [18].

Neural Network

Sentiment Analysis also makes use of the deep learning algorithms, neural networks.The latter is a cascade of neuron layers with output of one layer fed as input to the next successive layer. The information of each layer gets passed on and processed in the next layer to make the data more meaningful as it progresses. Since these models cannot process direct words, the use of preprocessing tools, embeddings or feature vectors are employed to convert the words into numerical forms [19].

- **Long short-term memory (LSTM)**: The algorithm LSTM is an advanced artificial neural network which has gained popularity in sentiment classification tasks. The model was first introduced in 1997 by Hochreiter and Schmidhuber [20] and since then other researchers have subsequently been enhancing and improving the model's accuracy [21].

- **Convolutional Neural Network**: CNN has gained its popularity for image processing, extracting the significant features of the image when the convolutional filter such as the kernel moves through the image. Hence, if the input data is given as one-dimensional this can be incorporated also for text classification [22]. Study shows that, applying convolutional neural network in sentiment analysis when properly trained can outperform other machine learning algorithms [23].

Hybrid Approach

Sentiment analysis has another method which is a combination of lexicon-based and machine learning algorithms. Also known as a hybrid approach. This type of model is extremely popular for SA, with sentiment lexicons playing a crucial role in the majority of systems. Even though this is a popular approach, few researchers propose a hybrid architecture involving both lexicon-based and machine learning techniques. Such hybrid techniques in sentiment analysis is utilising two machine learning algorithms [14].

2.2.5 Sentiment Analysis in Recommender System

Researcher Satvik Garg conducted a study in 2021 on a drug recommendation system based on sentiment analysis, utilising machine learning techniques. Various algorithms were used including: Logistic Regression, Ridge Classifier, Perceptron, LinearSVC, Multinomial Naïve Bayes, and Stochastic Gradient Descent, applied to Bag of Words (BOW) and TF-IDF. Furthermore, classifiers like Decision Tree, LGBM, Random Forest, and CatBoost used Word2Vec and Manual Features. The Drug Review dataset was used, taken from the UCI ML repository which was cleaned and preprocessed for converting text into numerical formats suitable for model training. Synthetic Minority Over-sampling Technique (SMOTE) was employed in the training data as was a class imbalance. The models were evaluated using precision, recall, F1-score, accuracy, and AUC score metrics. Resulting, LinearSVC with TF-IDF vecotrisation was the most effective, achieving 93% accuracy, while the Decision Tree with Word2Vec was the least effective.

Hence, the best-predicted values were taken from each method; 93% from LinearSVC on TF-IDF, 91% from LGBM on Word2Vec, 91% from Perceptron on BOW, and 88% from Random Forest on manual features. The values were multiplied by normalised useful counts to calculate the overall drug score by condition. Despite these performances, the recommendation system was still not ready for real-live applications [24].

Another research conducted by Bui Thanh Hung in 2020 explored a hybrid deep learning technique utilising a combination of CNN and LSTM. This combination was used for sentiment analysis and integrated in a recommender system. The Continuous Bag of Words (CBOW) model was employed for generating word embeddings to capture the semantic and syntactic nuances of words from the Amazon food reviews dataset. The CNN-LSTM model processes the prepared evaluations to classify sentiments as positive, neutral, or negative. Additionally, the results of the sentiment analysis were then incorporated with a user-item rating matrix to use a matrix factorisation technique, aiming to refine the prediction accuracy of user preferences. The CNN-LSTM model outperformed both the individual CNN and LSTM algorithms with an 83.45% sentiment analysis accuracy. Furthermore, this method had a reduced Root Mean Square Error (RMSE) when combined with matrix factorisation, having better predictive accuracy. As a result, sentiment analysis combined with a matrix factorisation model performs significantly better than the traditional recommender system [25].

Research on enhancing traditional collaborative filtering recommender systems by incorporating sentiment analysis was proposed by Ikram Karabila et al. It involved implementing a Bidirectional Long Short-Term Memory model, or also known as Bi-LSTM. The datasets used were the Kindle book review which consisted of a diverse range of book categories and the Amazon digital music dataset consisting of customer reviews on digital music products. For sentiment analysis, GloVe embeddings were used as the word representation method. In addition, several classification metrics, including accuracy, F1-score, and Area under the ROC Curve (AUC), were used to assess the model. Collaborative filtering with cosine similarity was employed both for the item-based and user-based approach to predict the rating that a user would rate on a particular product that has not been rated.

Sentiments were classified as positive or negative, after that incorporated as an additional feature for the recommender system. Hence, the model accurately predicted sentiments for both datasets. As a result, the model for sentiment analysis was incorporated into the recommender systems, and as table 2.3 and 2.4 [26] demonstrate, these systems perform better than the recommender system without sentiment analysis on the Kindle book dataset and the Amazon digital music dataset. This proposed study suggested that employing sentiment analysis on recommender systems leads to better performance.

Table 2.3: *User-based performance metrics.*

Dataset	MAE w/o SA	RMSE w/o SA	MAE w/SA ($\alpha = 0.3$)	RMSE w/SA ($\alpha = 0.3$)	MAE w/SA ($\alpha = 0.5$)	RMSE w/SA ($\alpha = 0.5$)	MAE w/SA ($\alpha = 0.7$)	RMSE w/SA ($\alpha = 0.7$)
Kindle Book Reviews	2.30	2.63	1.16	1.34	1.13	1.31	1.12	1.30
Amazon Digital Music	2.18	2.66	1.33	1.39	1.17	1.31	1.15	1.28

Table 2.4: *Item-based performance metrics.*

Dataset	MAE w/o SA	RMSE w/o SA	MAE w/SA ($\alpha = 0.3$)	RMSE w/SA ($\alpha = 0.3$)	MAE w/SA ($\alpha = 0.5$)	RMSE w/SA ($\alpha = 0.5$)	MAE w/SA ($\alpha = 0.7$)	RMSE w/SA ($\alpha = 0.7$)
Kindle Book Reviews	2.36	2.73	1.18	1.35	1.25	1.45	1.54	1.80
Amazon Digital Music	1.96	2.55	1.32	1.46	1.58	1.84	1.91	2.49

This chapter concludes with a survey of the literature on the analysis of e-commerce-based recommender systems. Additionally, the studies by [24] [25] [26] demonstrate the integration of sentiment analysis in recommender systems and highlight the adaptability and effectiveness of sentiment analysis in improving recommendation systems. However, each study employs a unique methodology that aligns with its research goals. According to [24], there is still room for improvement before the recommender system can be used in a real-world application. Whereas, other studies [25] [26] concluded that recommender systems that included sentiment analysis performed better than those without sentiment analysis. Therefore, their findings enrich this field and offer further investigation. The next chapter will cover the Research Methodology, encompassing the research strategy, data collection methods and tools, prototype, ethical considerations.

Chapter 3: Research Methodology

+ This section outlines the methodology employed for this research. It offers a thorough depiction of the evaluated information. Furthermore, this chapter follows a research pipeline for the development of the prototype. It presents the utilisation of machine learning techniques for sentiment analysis and recommender systems for an e-commerce dataset.

3.1 Defining the Problem

Recommender systems, despite their widespread use and gaining popularity in many areas, especially on e-commerce website have faced several challenges. Such as efficiently personalising product recommendations to the preferences of users. Even though significant advancement have been made to develop reliable RS, it is crucial to acknowledge and address the problems they come across. In addition, traditional recommender systems completely rely on ratings to suggest items to users. However, utilising this feature alone may not be sufficient and have a limited view on the users preferences.

3.2 Problem Solution

Thus, to overcome this limitation, incorporating additional features such as user reviews are considered to reduce this problem. Customer reviews provide extra information based on the fact that numerical ratings alone may fail to capture the users' preferences. Therefore, classifying and analysing reviews assist recommender systems to minimise the limitations. Thus, applying sentiment score on reviews whether they are positive or negative [26]. Taking this solution into consideration was the main motivation in choosing this area to study. For this reason, the aim of this project is to explore and determine how ML sentiment analysis can be integrated into an e-commerce recommender system to generate high-quality recommendations for a specific user.

Moreover, the study will investigate the traditional recommender systems. The hypothesis of this project is that applying sentiment analysis on traditional recommender systems as a feature improves the overall performance of the system to recommend to users based on their preferences. To answer the hypothesis the following research questions will be explored:

1. Which model performs better in terms of accuracy, precision, recall, F1-score, and AUC score?
2. Are deep learning algorithms better than traditional machine learning algorithms in terms of sentiment analysis?
3. Can sentiment analysis boost the performance of traditional recommender system?

3.3 Research Process

A quantitative research approach will be taken. Since secondary data is being utilised, no human interactions is involved, and no businesses or company names will be mentioned in this study. Therefore, there would be no ethical issues since the data is taken from a public source. This approach is employed as there are already existing datasets for our development to build an e-commerce sentiment analysis recommender system. Quantitative was chosen over qualitative methods, since the data provides a more structured and standardised approach to data collection analysis, allowing for the objective measurement of variables. Whereas, of the data interpretation qualitative methods is more subjective. Furthermore, quantitative methods are less time consuming and more cost-effective when using already existing datasets. For the development of the prototype, machine learning pipeline is used since this research is going to make use of machine learning algorithms. Figure 3.1 shows a machine learning pipeline that will be adapted for this research.

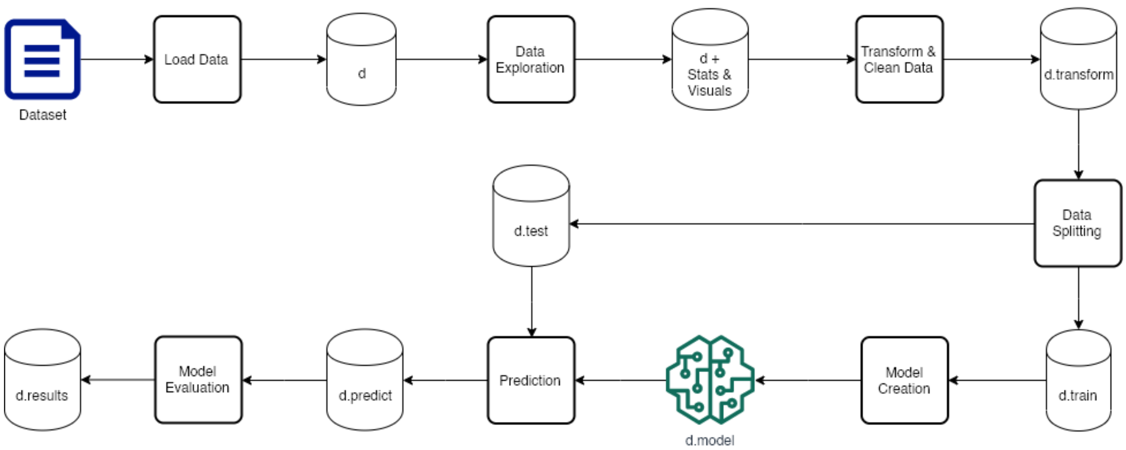

Figure 3.1: *Research Pipeline*

Since this study is based on an e-commerce recommender system the dataset that will be utilised is the Amazon Product. The Amazon Product dataset by [27] [28], which contains data spanning from May 1996 to July 2014. The total data consists of 142.8 million reviews which would be too large for our study. Therefore, this was narrowed down for a specific field, the Cell Phones and Accessories 5-core dataset. The 5-core dataset consists of 194,439 reviews that users gave to the Amazon products.

For the development of the prototype, the tools and languages that will be utilised include: Visual Studio Code with Jupyter Notebook; and Anaconda. In addition, the python language will be used with necessary libraries. A GPU environment was needed to be used since deep learning algorithms are computationally intensive. This includes installing the cudatoolkit, TensorFlow and Pytorch. The reason for this is due to the fact that deep learning algorithms are computationally intensive and would require a GPU environment. The NLTK will be used for the preprocessing of the reviews to be ready for the models. Additionally, for visualisation and cleaning of data, pandas and matplotlib will be utilised.

The proposed system of this study is the development of a recommender system utilising Sentiment Analysis. Consequently, a step-by-step of the machine learning pipeline will be discussed.

- **Data Loading**: The Review Cell Phone and Accessories dataset is gathered and loaded from the Amazon Review Product data.
- **Data Exploration**: The initial phase of data exploration was focused mainly with understanding the dataset's structure and quality. This evaluation included in calculating the number of rows, checking for null values, and identifying any duplicate entries. Furthermore, the lengths of the reviews were analysed, and the mean of these lengths was calculated to determine the average review size.
- **Data Visualisation and Data cleaning**: During the data visualisation and

cleaning phases, numerous tasks were carried out to visually present and interpret the dataset, which aided in identifying underlying patterns and insights. Initially, null values and any duplicate columns were removed. Attention was then focused on finding the distribution of review lengths to measure the average. The distribution of ratings was shown to define the most popular ratings given by users. For the sentiment analysis part, according to a study presented by Liu in 2012 documented that research in sentiment classification analysis do not employ the neutral class to simplify the classification problem [29]. Hence, taking into consideration this proposed research the sentiment analysis will be divided into positive (1) and negative (0) sentiments. Subsequently, the distribution of these binary sentiment values was displayed, showing a clear count of positive against negative sentiments. This visualisation not only offered an overview of the general sentiment distribution but also opened up more steps to focus on for the model to be successful.

- **Data Transformation**: During the data transformation phase, the ratings were converted into positive (1) and negative (0) sentiments, with ratings above three deemed positive and those below three considered negative. As previously discussed in the data cleaning stage, the neutral class was excluded to streamline the classification process. Additionally, the review column was filtered to match the average length of reviews. For preprocessing, the review column was modified. This included: converting text to lowercase, removing punctuation, special characters, and numbers, the use of

NLTK library for tokenisation, removal of stop words, and lemmatisation. In preparation for model development, feature extraction was conducted using the TF-IDF vectoriser and sequence padding was applied for the deep learning algorithms.

- **Data Splitting**: In the data splitting phase, both sentiment analysis and the recommender system were divided using 80% of the data for training and 20% for testing. This stage utilised two train-test split libraries. For the sentiment analysis the sklearn train-test split functionality was employed, ensuring that data was appropriately divided into train and validate the performance of the machine learning models. For the recommender system, which utilises the SVD model, the Surprise library needed to be utilised. In addition, as concerns developed with the data being unbalanced, oversampling methods were introduced to achieve a balanced distribution of sentiment classes. Since the data may substantially impact the models performance, particularly in classification tasks,

- **Model Creation**: In the model creation phase, the project is split into two parts, sentiment analysis and recommender system. For the sentiment analysis, the utilisation of five state-of-the-art machine learning algorithms including: Logistic Regression, Support Vector Machine (SVM), Random Forest, Long Short-Term Memory (LSTM); and Convolutional Neural Network (CNN). The GridSearchCV was employed for cross validation and hyperparameter tuning of the traditional machine algorithms, whereas the Keras Tuner to fine-tune the hyperparameters of the deep learning models, namely

LSTM and CNN. The deep learning algorithms had to be optimised, hence the Adam optimiser was used. For the second part of our project, the recommender system adopted a collaborative filtering approach employing a Singular Value Decomposition (SVD) model. This technique is well-suited for making predictions about user preferences based on past interactions within the user-item matrix. Moreover, there was a development of three distinct types of recommendation systems to evaluate the impact of incorporating sentiment analysis into the process:

1. **Basic Recommendation System**: This method uses only the ratings feature and serves as our baseline model.
2. **Sentiment Recommendation System**: This technique integrates sentiment analysis, and is utilised as the only feature of the system.
3. **Combined Recommendation System**: This model then utilises both ratings and sentiment analysis to provide recommendations that considers both the rating score and the sentiment score of the review.

This planned strategy allows for the evaluation of the performance of each model and the assessment of how sentiment analysis might improve the customisation and accuracy of product recommendations.

- **Prediction**: In this stage of the research pipeline, the machine learning that performs the best was used to predict sentiment scores for every review. The best model classified reviews either positive or negative sentiments. Additionally, the sentiment model was employed into the recommendation system utilising a Collaborative filtering, SVD model. The combined recommender system was then used to suggest the top 5 items that a user might like.

- **Model Evaluation**: For this research study both the sentiment analysis and recommender system needed to be evaluated. For the sentiment analysis models the metrics used include:

  1. **Precision**: This metric measure the accuracy of positive predictions.
  2. **Accuracy**: Gives the percentage of total correct predictions, both positive and negative, which offers a general sense of model effectiveness on the data.
  3. **Recall**: Indicates the ability of the model to identify and find all the relevant instances in the dataset.
  4. **F1-score**: F1-score is a combination of precision and recall into a single metric by taking the harmonic mean.
  5. **AUC score**: Area Under Curve score assesses the model's ability to distinguish between classes at various thresholds. This metric is vital for understanding the trade offs between True Positive (TP) and False Positive (FP).

  In terms of the recommender system, the metrics used to evaluate the accuracy of predictions include:

  1. **Root Mean Square Error (RMSE)**: Measures the average of the errors between predicted and actual ratings. It gives a sense of how far

the predictions are from the actual values, with lower values indicating better performance.

2. **Mean Absolute Error (MAE)**: Calculates the average absolute difference between predicted and actual ratings.

This current research study was adapted to explore the integration of machine learning SA in recommender system, an e-commerce approach. The objective of this study was to enhance the accuracy and prediction of recommending items to users, utilising classification models for sentiment analysis. Furthermore, applying a collaborative filtering approach with an SVD model for recommender systems.

The development of this prototype was methodically followed using a machine learning pipeline. Beginning with data loading, data exploration, visualisation, data cleaning, data transformation, data splitting, model creation, prediction, and model evaluation. Every stage of the research pipeline was taken into consideration to address and answer the research questions.

Hence, the next chapter, Analysis of Results and Discussion, will investigate the findings found during development. It will discuss the outcomes of the sentiment analysis models, and when applied into the recommender system. This evaluation is crucial to understand the practical applications and limitations of this study. It sets the stage for future recommendations and innovations in this research area.

Chapter 4: Analysis of Results and Discussion

4.1 Findings

This section will go through the findings that were found during each stage of the research methodology. Consequently, when the data was loaded into the Python environment using the Pandas library, the data was of JSON format. Once the data was loaded the exploration stage was conducted. This resulted in the dataset consisting of 194,439 rows and 9 columns. Additionally, the data cleaning and transformation tasks were described in more detail in the research methodology chapter. Visuals were created to examine the data and identify any threats that could be a problem to the models.

Figure 4.1 shows the distribution of the review lengths, determining that most of them were in the range of 0 to 100. Therefore, filtering was required to aid the model.

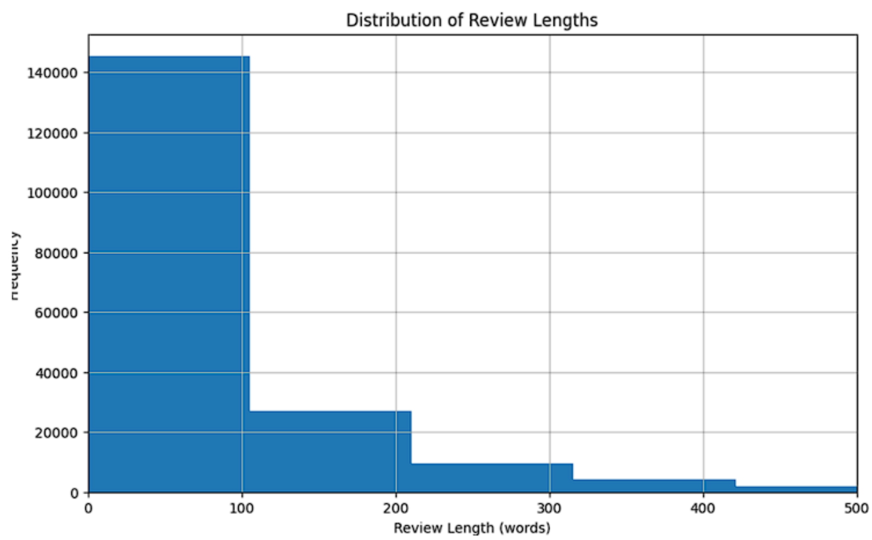

Figure 4.1: *Distribution of Review Lengths*

Consequently, the next figure 4.2 shows the distribution of ratings ranging from 1 to 5 where 5 is the most positive and 1 is the most negative. This visual indicates that the majority of the ratings are positive. This means that users are more likely to give a higher rating.

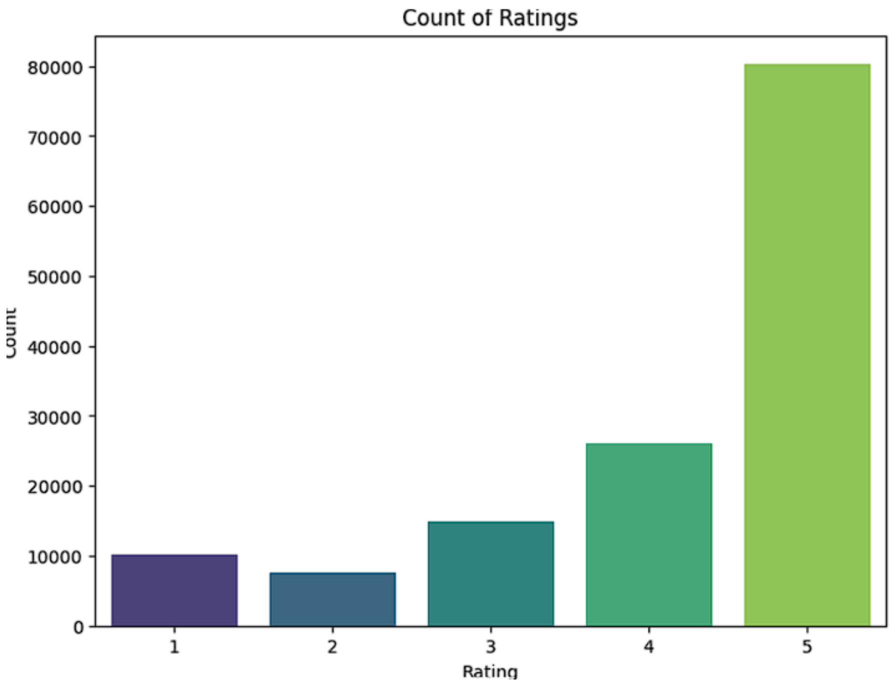

Figure 4.2: *Count of Ratings*

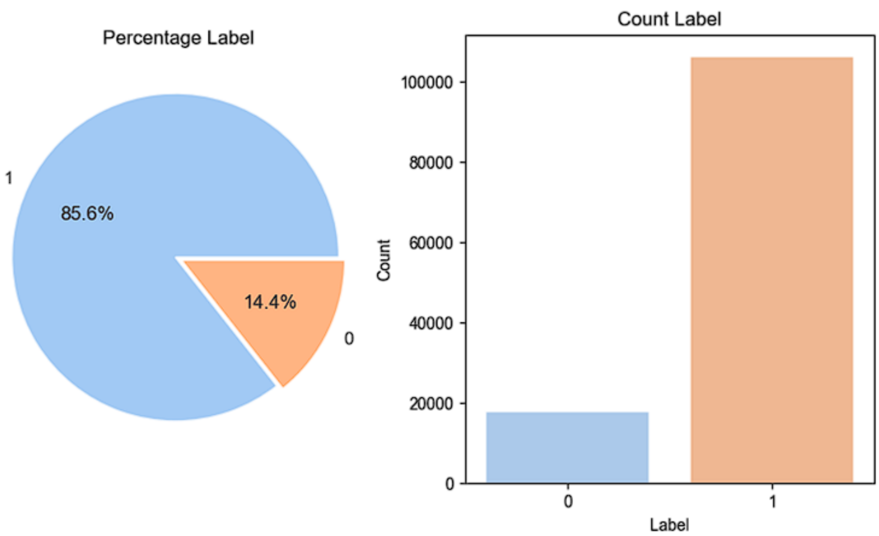

Figure 4.3: *Positive and Negative Sentiments*

Figure 4.3 shows the distribution of sentiments where the positive sentiment labeled as 1 has 85.6% of the dataset, whereas the negative sentiment has 14.4%. The visuals 4.2 and 4.3 can conclude that there is an imbalanced data since, most of the data consists of having positive ratings and sentiments. Therefore, an oversampling technique was considered when training and testing the sentiment models.

The dataset was split into training and testing sets, with a 70% allocated for the training and 30% for testing. Traditional machine learning and deep learning algorithms were used for sentiment analysis. This included Logistic Regression, SVM, Random Forest, CNN, and LSTM. The performance of these techniques were evaluated using several evaluation metrics. The metrics used for sentiment analysis include: Accuracy, Precision, Recall, F1-score, and AUC score. Table 4.1 shows the results of the sentiment models.

	Accuracy	Precision	Recall	F1-score	AUC score
Logistic Regression	0.89	0.90	0.89	0.89	0.92
Random Forest	0.91	0.91	0.91	0.89	0.94
SVM	0.93	0.93	0.93	0.92	0.95
CNN	0.92	0.91	0.92	0.91	0.93
LSTM	0.90	0.90	0.90	0.90	0.89

Table 4.1: *Sentiment Analysis Results*

- **Logistic Regression**: The performance of this model resulted in having an Accuracy, Recall, and F1-score of 0.89, Precision of 0.90, and 0.92 in AUC score.

- **Random Forest**: Random Forest achieved 0.91 in accuracy, Precision and Recall; 0.89 in F1-score and 0.94 in AUC score.

- **SVM**: The Support Vector Machine algorithm obtained an 0.93 in Accuracy, Precision and Recall; a 0.92 in F1-score and an 0.95 in AUC score.
- **CNN**: This deep learning technique achieved a 0.92 score in Accuracy and Recall; 0.91 in Precision and F1-score; and 0.93 in AUC score.
- **LSTM**: The Long Short term Memory model performed a 0.90 in Accuracy, Precision, Recall, and F1-score, and a 0.89 in AUC.

The overall performance of each classification model for sentiment analysis performed very well. In addition, table 4.1 concludes and answers the first two research questions.

The second part of the study was integrating sentiment analysis with recommender systems, which will answer the last research question. The results of the three recommender systems that were built for this study is shown in table 4.2. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were utilised to identify which recommender system performs better. The baseline recommender system had an RMSE of 1.20 and MAE of 0.87. Whereas, the Sentiment Analysis recommender system achieved an RMSE of 0.43 and 0.42 MAE. In addition, the combined model resulted in having an RMSE of 1.14 and MAE of 0.79.

	RMSE	MAE
Baseline Recommender System	1.20	0.87
Sentiment Analysis Recommender System	0.43	0.42
Combined Recommender System	1.14	0.79

Table 4.2: *Recommender Systems Results*

Additionally, the combined recommender system that uses both the sentiment analysis feature and ratings of users, was then utilised to recommend items for a specific user. Figure 4.4 shows the top 5 recommendations for user A30TL5EWN6DFXT.

```
Top 5 recommendations for A30TL5EWN6DFXT:
Item: B006MK0ZD4, Estimated Rating: 4.788100841126396
Item: B00AAKGF72, Estimated Rating: 4.755232320008208
Item: B006JHU3DC, Estimated Rating: 4.752616409347736
Item: B0085QRJ5U, Estimated Rating: 4.747266224773453
Item: B00CE3IC74, Estimated Rating: 4.696636605652976
```

Figure 4.4: *Top 5 Recommendations*

4.2 Compare with Existing Research

The research conducted by Satvik Garg's in 2021 [24] which is documented in Chapter 2 provided a detailed comparison of various machine learning models for sentiment analysis. Particularly, highlighting the LinearSVC model, which had an impressive score of 93% accuracy when used with TF-IDF vectorisation. This resulted in establishing and opening up the field of classification tasks for sentiment analysis. Similarly, with Satvik Garg's research this project employs the TF-IDF vectorisation technique. This ultimately, prepares the reviews for machine learning by emphasising the significance of words in a document.

The model that stood out the most from the other machine learning algorithms which were used, was the SVM, achieving a 93% accuracy in testing. The SVM adopted a RBF kernel and a regularisation parameter (C) set at 10. This result is significant when compared with Garg's findings as it reaffirms the effectiveness and adaptability of SVMs in handling text classification tasks.

Studies conducted by Bui Thanh Hung, in 2020 [25] and Ikram Karabila, in 2023 [26] explore hybrid deep learning techniques and the integration of sentiment analysis in recommender systems. The researchers accomplished impressive results in accuracy and predictive performance. Bui's research, employed a CNN-LSTM model with a CBOW approach for word embeddings, utilising the Amazon food reviews. Effectively classifying sentiments and enhancing user-item matrix predictions through matrix factorisation. The approach used significantly reduced the RMSE score and improved the sentiment accuracy to 83.45%, which surpasses the individual algorithms CNN and LSTM.

Similarly, Karabila employed a Bi-LSTM with GloVe embeddings on the Kindle book reviews dataset and the Amazon digital music dataset. This study also integrated sentiment analysis into collaborative filtering recommender systems. This research results show an enhanced performance in lower RMSE and MAE when applying sentiment analysis as an additional feature. The results presented in table 2.3 and table 2.4, display the improved performance across varying levels of sentiment analysis integration.

A comparison of these studies with the current project, utilising a collaborative SVD recommender system and employing a SVM model for sentiment analysis. An advancement can be seen in the results of the recommender systems. This current study focused on the development of three recommender systems; the baseline, sentiment analysis recommender system and a combination of sentiment analysis and ratings. When sentiment analysis was applied, the system demonstrated a better predictive accuracy, with the sentiment analysis recommender system achieving the lowest RMSE of 0.43 and MAE of 0.42. Hence, these results are considerably better than those referenced in the Literature Review. Moreover, this highlights the effectiveness of integrating sentiment analysis directly into the recommender system. In addition, different studies consistently show an improvement in the prediction accuracy of recommender system.

4.3 Addressing Research Questions

4.3.1 Research Question 1: Model Performance

The first research question states which model performs better in terms of accuracy, precision, recall, F1-score and AUC score. On that account, the results shown in table 1, conclude that the Support Vector Machine emerged as the top performer. This is due to adjusting the model using the radial basis function (RBF) kernel and an optimal regularisation parameter (C) of 10. The SVM algorithm demonstrated superior capabilities in handling classification task problems. It achieved impressive scores with an Accuracy, Precision, and Recall of 0.93, 0.92 in F1-score and 0.95 in AUC score. These metrics point out the robustness and precision of the models, making it particularly effective in distinguishing between the classes in high dimensional spaces.

4.3.2 Research Question 2: Traditional Machine Learning vs Deep Learning

Traditional machine learning algorithms that were used for the project were then compared with deep learning algorithms, this was necessary to answer the second research question. Even though the algorithms all performed very well, there was one model that outperformed. In the case of the LSTM model, it showed the least overall performance, particularly in the AUC score of a 0.89 among all techniques. A suggestion to this performance would be that while LSTM is well-suited for sequence prediction tasks, it may require further tuning and possibly a more complex network architecture. However, the CNN model performed the second best out of all the algorithms. Nonetheless, the SVM model stood out with its Accuracy, Precision, and Recall at 0.93, F1-score at 0.92, and AUC score at 0.95. Similarly, Random Forest demonstrated a strong performance. Closely following SVM, especially noted in its AUC score. To enhance their capabilities for sentiment analysis tasks, the CNN and LSTM models require further investigation, within their network architectures and training data.

These observations suggest that for the given dataset and the problem statement, traditional machine learning techniques, particularly SVM, provided better performance when compared with deep learning models. This indicates that applying more classical and simpler approaches may still be preferable depending on the complexity of the dataset.

4.3.3 Research Question 3: Recommender Systems Performance

The incorporation of sentiment analysis was applied into the recommender systems to check whether it enhances the performance of the systems. Consequently, answering the third research question of this study. Table 4.2 establishes that the recommender system, which only uses the sentiment analysis as a feature, had the lowest RMSE of 0.43 and MAE of 0.42. This system significantly outperformed the baseline recommender system, having a RMSE and MAE values of 1.20 and 0.87. This marked improvement suggests that sentiment analysis substantially boosts the accuracy and predictive quality of the recommendations. Additionally, when SA was combined with the ratings feature, it also improved the performance of the model, having a RMSE and MAE values of 1.14 and 0.79.

These results indicate that both the sentiment analysis model and the combined model offer enhancements over the baseline model. Therefore, applying SA will improve the prediction and refining of what the user might like in terms of textual reviews. This evidence strongly supports the incorporation of sentiment analysis into traditional recommendation algorithms, showcasing its potential to provide an understanding of user preferences and improve quality of recommendations.

4.4 Addressing Hypothesis

The hypothesis of this research is that when applying sentiment analysis on recommender systems as a feature, it improves the overall performance of the system to recommend to users based on their preferences.

After a thorough investigation in addressing the research questions of this project, it is evident that the hypothesis confirmed to be true. Incorporating sentiment analysis into recommender systems significantly enhances their performance. The findings established an improvement in accuracy and predictive quality. In particular, the sentiment analysis recommender system outperformed the baseline model with lower RMSE and MAE value. This validates that sentiment analysis effectively captures user preferences, leading to more tailored and accurate recommendations. Hence, upholding the hypothesis that when applying SA in RS it does in actual fact improve the performance of the system.

Moving on to the next chapter, the Conclusion, will provide a comprehensive summary of the entire research paper. This chapter will include: a review of the key findings; the overall contributions of the research; and suggestions for future work and limitations that was found in the process of this study.

Chapter 5: Conclusions and Recommendations

This chapter will provide a summary of the research that was undertaken on enhancing e-commerce recommender systems when implementing sentiment analysis. It includes: a summary of the key results; an evaluation of the research questions addressed; future recommendations; and a discussion of the limitations encountered during the development of the prototype. Lastly, revisiting the initial hypothesis to confirm its validity based on the research findings and results.

5.1 Summary of Results

Recommender systems have been a highly sought research area, as several scholars have created and improved various models to accurately predict recommendations to users. In addition, this research area connected to other topics such as the integration of sentiment analysis into recommender systems. As, the primary goal of recommender systems is to give accurate and personalised recommendations to its users. Furthermore, this research adopted the use of sentiment analysis on RS utilising the Cell Phone and Accessories 5-core dataset. An evaluation of five models for SA was used, including Logistic Regression, Random Forest, SVM, CNN, and LSTM.

The SVM model resulted to be the top performer across multiple evaluation metrics. This model demonstrated to have superior capabilities in handling classification tasks. It achieved an accuracy of 0.93, precision of 0.93, recall of 0.93, F1-score of 0.92 and an AUC score of 0.95. This establishes that ML algorithms provided a better performance when compared to deep learning models (CNN and LSTM) for this dataset and problem statement. Consequently, the best model was applied for sentiment analysis and was integrated with a recommender system to compare with two other systems, the baseline and the combined techniques.

The RS models employed a collaborative filtering approach using an SVD algorithm. The baseline used only the user ratings, whilst the sentiment model used the sentiment feature. Whereas, the last recommender combined both sentiment and ratings together. This resulted in the SA approach achieving the lowest RMSE and MAE of 0.43 and 0.42. The combined system achieved an RMSE of 1.20 and an MAE of 0.87. When applying sentiment analysis in both models as a feature, better scores were established when compared with the baseline method. Therefore, these findings capture the effectiveness of integrating sentiment analysis into recommender systems.

From the comprehensive analysis conducted in answering the research questions for this study, it is evident that the hypothesis results to be true. This is shown from the improvements of RMSE and MAE in the recommender systems. For this reason, applying sentiment analysis into recommender systems enhances the accuracy and prediction of capturing the user preferences. Thus, validating the hypothesis.

5.2 Future Work Recommendations

Future research could explore several promising areas to improve the performance of sentiment analysis into recommender system. This include:

5.2.1 Applying Other models For Sentiment Analysis

Researchers could investigate on utilising more models and could even incorporate a hybrid solution for sentiment classification tasks. Such models include CatBoost and XGBoost for traditional machine learning algorithms and Recurrent Neural Network for deep learning.

5.2.2 Enhancing Recommender Systems

Different algorithms can be applied rather then using traditional algorithms such as the Collaborative filtering approach. Implementing deep learning techniques in recommender system and even hybrid method can assist in capturing more complex user-item interactions to provide more personalised recommendations.

5.2.3 Exploring Different Datasets

The methodology was aimed at exploring the Cell Phone and Accessories 5-core dataset, an e-commerce approach. Moreover, researchers can explore in other research areas such as movies or books with varying characteristics to ensure the robustness and versatility of the findings.

5.2.4 Addressing Potential Challenges

Future studies should also address potential challenges found in recommender systems such as the Cold-Start Problem, Data Sparsity. Techniques that can be explored include domain adaptation, or incorporating user demographics or item attributes to mitigate these challenges.

5.3 Limitations

The investigation was conducted on a computer having an Intel i7 1170F 11th generation processor, 32GB of RAM and Gigabyte NVIDIA RTX 3060. During the models training process, the execution time of some of the models exceeded 60 minutes. Additionally, the CPU temperature reached 100 degrees Celsius, which poses a risk of overheating and potential damage. Hence, addressing these computational challenges will be crucial in future work to ensure more efficient model training and evaluation.

5.4 Conclusion

In summary, this paper successfully validate the potential of applying sentiment analysis on traditional recommender systems as a feature improves the overall performance of the system to recommend to users based on their preferences. Therefore, making this study a valuable insight to the field.

List of References

[1] Anjali Gupta. E-commerce: Role of e-commerce in today's business. *International Journal of Computing and Corporate Research*, 4(1):1–8, 2014.

[2] F Furtado and A Singh. Movie recommendation system using machine learning. *International journal of research in industrial engineering*, 9(1):84–98, 2020.

[3] Meenu Gupta, Aditya Thakkar, Vishal Gupta, Dhruv Pratap Singh Rathore, et al. Movie recommender system using collaborative filtering. In *2020 international conference on electronics and sustainable communication systems (ICESC)*, pages 415–420. IEEE, 2020.

[4] Sanjeevan Sivapalan, Alireza Sadeghian, Hossein Rahnama, and Asad M Madni. Recommender systems in e-commerce. In *2014 World Automation Congress (WAC)*, pages 179–184. IEEE, 2014.

[5] J Ben Schafer, Joseph Konstan, and John Riedl. Recommender systems in e-commerce. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 158–166, 1999.

[6] Xuesong Zhao. A study on e-commerce recommender system based on big data. In *2019 IEEE 4th international conference on cloud computing and big data analysis (ICCCBDA)*, pages 222–226. IEEE, 2019.

[7] Folasade Olubusola Isinkaye, Yetunde O Folajimi, and Bolande Adefowoke Ojokoh. Recommendation systems: Principles, methods and evaluation. *Egyptian informatics journal*, 16(3):261–273, 2015.

[8] PH Aditya, Indra Budi, and Qorib Munajat. A comparative analysis of memory-based and model-based collaborative filtering on the implementation of recommender system for e-commerce in indonesia: A case study pt x. In *2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pages 303–308. IEEE, 2016.

[9] Liaoliang Jiang, Yuting Cheng, Li Yang, Jing Li, Hongyang Yan, and Xiaoqin Wang. A trust-based collaborative filtering algorithm for e-commerce recommendation system. *Journal of ambient intelligence and humanized computing*, 10:3023–3034, 2019.

[10] Mahesh Goyani and Neha Chaurasiya. A review of movie recommendation system: Limitations, survey and challenges. *ELCVIA: electronic letters on computer vision and image analysis*, 19(3):0018–37, 2020.

[11] Ari Nurcahya and Supriyanto Supriyanto. Content-based recommender system architecture for similar e-commerce products. *Jurnal Informatika Ahmad Dahlan*, 14(3):90–101, 2020.

[12] R Sabitha, S Vaishnavi, S Karthik, and RM Bhavadharini. User interaction based recommender system using machine learning. *Intelligent Automation & Soft Computing*, 31(2), 2022.

[13] Nurul Aida Osman, Shahrul Azman Mohd Noah, Mohammad Darwich, and Masnizah Mohd. Integrating contextual sentiment analysis in collaborative recommender systems. *Plos one*, 16(3):e0248695, 2021.

[14] Mayur Wankhade, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780, 2022.

[15] Cach N Dang, María N Moreno-García, and Fernando De la Prieta. An approach to integrating sentiment analysis into recommender systems. *Sensors*, 21(16):5666, 2021.

[16] María del Pilar Salas-Zárate, Jose Medina-Moreira, Katty Lagos-Ortiz, Harry Luna-Aveiga, Miguel Angel Rodriguez-Garcia, Rafael Valencia-Garcia, et al. Sentiment analysis on tweets about diabetes: an aspect-level approach. *Computational and mathematical methods in medicine*, 2017, 2017.

[17] Anton Borg and Martin Boldt. Using vader sentiment and svm for predicting customer response sentiment. *Expert Systems with Applications*, 162:113746, 2020.

[18] M Ali Fauzi. Random forest approach fo sentiment analysis in indonesian. *Indones. J. Electr. Eng. Comput. Sci*, 12:46–50, 2018.

[19] Sujata Rani and Parteek Kumar. Deep learning based sentiment analysis using convolution neural network. *Arabian Journal for Science and Engineering*, 44:3305–3314, 2019.

[20] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[21] GSN Murthy, Shanmukha Rao Allu, Bhargavi Andhavarapu, Mounika Bagadi, and Mounika Belusonti. Text based sentiment analysis using lstm. *Int. J. Eng. Res. Tech. Res*, 9(05), 2020.

[22] Hannah Kim and Young-Seob Jeong. Sentiment classification using convolutional neural networks. *Applied Sciences*, 9(11):2347, 2019.

[23] Srikanth Tammina and Sunil Annareddy. Sentiment analysis on customer reviews using convolutional neural network. In *2020 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–6. IEEE, 2020.

[24] Satvik Garg. Drug recommendation system based on sentiment analysis of drug reviews using machine learning. In *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 175–181. IEEE, 2021.

[25] Bui Thanh Hung. Integrating sentiment analysis in recommender systems. *Reliability and Statistical Computing: Modeling, Methods and Applications*, pages 127–137, 2020.

[26] Ikram Karabila, Nossayba Darraz, Anas El-Ansari, Nabil Alami, and Mostafa El Mallahi. Enhancing collaborative filtering-based recommender system using sentiment analysis. *Future Internet*, 15(7):235, 2023.

[27] Ruining He and Julian McAuley. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *proceedings of the 25th international conference on world wide web*, pages 507–517, 2016.

[28] Julian McAuley, Christopher Targett, Qinfeng Shi, and Anton Van Den Hengel. Image-based recommendations on styles and substitutes. In *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval*, pages 43–52, 2015.

[29] Bing Liu. Sentiment analysis and opinion mining. volume 5, 05 2012.

Chapter A: Sample Code

For further details and to access the code related to the Sentiment Analysis Recommender System used in this study, please visit the following GitHub repository: Sentiment Analysis Recommender System.