FACULTY OF SCIENCE & TECHNOLOGY

MSc Data Science and Artificial Intelligence

May 2023

Fake Product Review Monitoring System

by

Shilpa Shaji Nellikakunnel

Faculty of Science & Technology
Department of Computing and Informatics
Individual Masters Project

Abstract

With development in e-commerce, online shopping takes a significant ratio of the whole market. The reviews are written by people who have bought the product and it has precise details about the products, which can be used as a guide while reviewing potential purchases. The quantity of users, nevertheless, has a significant impact on the review's fineness. Also, the users choose to write brief comments that describe the product's quality and cost. Sellers will occasionally use faked reviews, frequently generated by bots, to raise ratings on their products on e-commerce platforms like Amazon. This can make customers less satisfied with their purchases. The importance of identifying spammers and faked reviews increases as the damages caused increases.

The main idea of this project revolves around analysing customer's reviews provided in online shopping sites, by proposing a hybrid deep learning model CNN-BiLSTM and conduct a study to make the prominent impact in performance of hybrids models when compared with stand alone models. Further, this project will be useful to identify and remove the fake reviews and provide actual reviews along with ratings related to the products for genuine purchases.

Dissertation Declaration

I agree that, should the University wish to retain it for reference purposes, a copy of my dissertation may be held by Bournemouth University normally for a period of 3 academic years. I understand that once the retention period has expired my dissertation will be destroyed.

Confidentiality

I confirm that this dissertation does not contain information of a commercial or confidential nature or include personal information other than that which would normally be in the public domain unless the relevant permissions have been obtained. Any information which identifies a particular individual's religious or political beliefs, information relating to their health, ethnicity, criminal history, or sex life has been anonymised unless permission has been granted for its publication from the person to whom it relates.

Copyright

The copyright for this dissertation remains with me.

Requests for Information

I agree that this dissertation may be made available as the result of a request for information under the Freedom of Information Act.

Signed: Shilpa

Name: Shilpa Shaji Nellikakunnel

Date: 25 May, 2023

Programme: Msc Data Science and Artificial Intelligence

Original Work Declaration

This dissertation and the project that it is based on are my own work, except where stated, in accordance with University regulations.

Signed: Shilpa

Name: Shilpa Shaji Nellikakunnel

Date: 25 May, 2023

Acknowledgements

I would like to express my sincere gratitude to all those who have contributed to the successful completion of my Master's project and the writing of this report.

First and foremost, I would like to extend my heartfelt appreciation to my supervisor, Dr Hari M Pandey, for his invaluable guidance, support, and encouragement throughout this project. His expertise, knowledge, and feedback have been instrumental in shaping the direction of my research and improving the quality of this report.

I am also grateful to the faculty and staff at Bournemouth University, particularly the Computing and Informatics, for providing me with the necessary resources, facilities, and academic environment to conduct my research effectively. Their dedication to fostering a conducive learning atmosphere has greatly contributed to my personal and academic growth.

I would like to thank my family and friends for their unwavering support and encouragement throughout my academic journey. Their love, understanding, and motivation have been a constant source of strength, pushing me to strive for excellence in my work.

Lastly, I would like to acknowledge the contributions of all the authors whose research and publications I have referenced in this report. Their groundbreaking work has served as a foundation for my own research and has enriched the overall quality of this project.

Although it is not possible to mention everyone individually, I want to extend my appreciation to all those who have supported me in one way or another during this project.

Thank you all for being an integral part of my academic journey and for your invaluable contributions.

Shilpa Shaji Nellikakunnel

Contents

List of Figures

List of Tables

1 Introduction

In the era of online shopping, reviews play a crucial role in aiding customers to make informed decisions about purchasing products. They provide a platform for customers to express their opinions and experiences with the product, which can influence other potential buyers. Prior to reviews, consumers frequently made purchases based on advertisements and promotional materials from businesses as well as personal suggestions from friends and family. As a result, people had limited access to the thoughts and impressions of other customers who had used the good or service they were considering. Online reviews have grown in popularity in recent years, becoming a vital source of knowledge for shoppers making choices (Nagamma et al. (2015)). However, the issue of fake reviews has gotten worse as online reviews have grown in popularity. Reviews that are fabricated in order to either help or hurt a business, product, or service are known as fake reviews. They may be used to advertise fake products or services, damage a business' internet reputation, and mislead customers. As a consequence, the identification of fake reviews has become an important research topic, in order to help consumers make informed purchasing choices (Chen and Li (2022)).

Overall, reviews are a valuable form of feedback that provide benefits to consumers, companies, and the wider online community. As such, they are likely to continue to play a significant role in the digital landscape for years to come. However, the authenticity of online reviews has become a significant concern, as fake reviews can skew the overall rating of a product and mislead customers.

1.1 Background

Prior to the advent of internet platforms such as e-commerce websites and social media, reviews were typically disseminated by traditional means of communication such as word-of-mouth, printed publications, or personal referrals. Reviews had a tiny audience and were mostly shared inside a small community or through offline media. Reviews then had a limited reach and people relied on personal interactions, discussions with friends, or recommendations from trusted sources to gather information about products, services, or experiences. Reviews often had a regional or localised impact, with suggestions and comments spreading within certain areas or communities. This made it difficult for consumers to access an extensive amount of feedback for items or services from various regions. The transition to internet platforms has greatly broadened the reach of reviews, allowing consumers to access a diverse range of perspectives, evaluations, and suggestions from a global community.

People are now able to share their experiences with a variety of products or services using text comments, also known as reviews, thanks to the shift from conventional marketing to online marketing. Consumers who frequently make purchases through e-commerce websites such as Amazon commonly publish reviews on the website. As discussed by Al-khiza'ay et al. (2019), Consumers are permitted to provide reviews that reflect their feelings about the items or services purchased. Product reviews have evolved into a significant form of customer-generated content that acts as a reputable information source. The data posted can be useful to a wide range of entities, including product designers, producers, potential buyers, and e-commerce website owners.AlZu'bi et al. (2019) mentioned that people can now more easily reach websites with reviews of all kinds of products, even rare ones. Before making a purchase, reviews serve as supplemental information to aid consumers in making appropriate choices. And hence buyers depend on these reviews and the reviewers.

However it is important for consumers to be cautious when reading reviews and to not rely solely on the overall rating, but also to read through individual reviews and to consider multiple sources. Fake reviews can harm the reputation of a company, lead to lost sales, and mislead consumers into buying a product that does not meet their expectations. Some common motivations for creating fake reviews are to boost the rating of a product, to harm the rating of a competitor's product, or to make money by selling fake reviews (Halim et al. (2022)).

The use of deep learning techniques for fake review detection is a critical research area, with potential applications in various industries. However, several challenges need to be addressed to develop accurate fake review detection systems. One of these challenges is class imbalance, where the number of fake reviews is much smaller than genuine reviews. Another issue is the lack of a comprehensive understanding of the characteristics of fake reviews. To overcome these challenges, previous studies have employed various methods, such as using different evaluation metrics, re-sampling techniques, adjusting class weights, and incorporating clustering techniques. In addition, several studies have conducted experiments to evaluate the effectiveness of different machine learning and deep learning models for fake review detection (Santhosh Krishna et al. (2022)).

1.2 Rationale and Motivation

The rationale and motivation for a fake product review monitoring system stem from the prevalence of fake reviews in online marketplaces and the negative impact they can have on both consumers and businesses. Nowadays on E-commerce sites, anyone can publish an opinion review on any product. If these reviews are intended to devalue any product or to offer a positive rating to a worthless product, they can constitute a threat. This can severely damage an organization's reputation. As a result, effective strategies and solutions to the mentioned challenge are required.

As mentioned by Santhosh Krishna et al. (2022), fake reviews can mislead consumers into purchasing products that do not meet their expectations, leading to lost sales and potential harm to the reputation of a business. In addition, fake reviews can harm the credibility of online marketplaces and discourage consumers from using them. Furthermore, this can lead to higher product returns, unfavourable word-of-mouth, and a reduction in client loyalty. The use of fake reviews can create an unfair advantage for businesses that engage in this practice, giving them an unfair competitive edge over honest businesses that rely on genuine customer feedback.

As a result, the development of a fake product review monitoring system is important to ensure that online marketplaces maintain their credibility and that consumers have access to honest and reliable information when making purchasing decisions. Such a system could help businesses identify and remove fake reviews and take appropriate actions against those who engage in this practice, ultimately benefiting both consumers and honest businesses.

1.3 Aims and Objectives

This project aims to develop a "Fake Product Review Monitoring System", and to accomplish this goal, six specific objectives have been established. These objectives include:

- Develop hybrid deep learning model where Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) which can to detect complicated textual patterns, helps to detect fake reviews and improve accuracy.

- Address class imbalance issues in the data

- Evaluate the performance of the proposed model on a publicly available data-set
- Compare the proposed model to other state-of-the-art machine or deep learning models
- Improve the quality of customer reviews in various industries
- Develop an accurate and efficient fake review detection system.

1.4 Research Questions and Objectives

- What is the proposed hybrid deep learning model techniques to detect fake reviews?

  The study aims to examine hybrid deep learning model and assessing their effectiveness and efficiency and then evaluate how well these technique work in practice.

- How does the proposed model address the issue of class imbalance in fake review detection?

  Studying and applying methods to address the issue of class imbalance in a data-set assessing their effectiveness and efficiency.

- How does the proposed model perform on a publicly available data-set in comparison to other state-of-the-art machine or deep learning models?

  Evaluate the model's performance on a publicly available data-set using established evaluation metrics.

- What is the accuracy and efficiency of the proposed fake review detection system?

  Critically examine past researchs, including identifying and understanding any limitations or shortcomings of previous studies. Develop a plan to address these issues which is more nuanced and well-informed, and design a study that is more likely to produce meaningful and impactful results.

- In what industries can the proposed system be used to improve the quality of customer reviews?

  Fake review detection can be used in various industries to improve the quality of customer reviews. Some examples of industries where this system could be used include e-commerce, hospitality, food and beverage, and tourism

1.5 Proposed Solution

This research aims to contribute to the development of effective hybrid deep learning model to identify and manage fake reviews in e-commerce sites. By providing recommendations to e-commerce companies, it helps them to improve the accuracy of false review identification, increase trust and credibility, give reliable product suggestions, improve user experience, decrease reputation threats, and assure regulatory compliance by applying this methodology. These advantages contribute to the growth and success of e-commerce businesses in a competitive online marketplace.

This study will use CNN and RNN models to classify the reviews as either authentic or fake. To evaluate the effectiveness of the proposed model, the study compares the proposed model with existing methods or techniques such as Random Forest, Naive Bayes, CNN and LSTM. The study uses stimulation metrics such as accuracy, precision, recall, and F1-score to measure the performance of different models to determine their efficiency, and conducts a comparitive study with the proposed model.

1.6 Thesis Organisation

The thesis comprises four chapters, each with its own focus and purpose. Here's a summary of what you can expect to find in each one.

- **Chapter 2** : The literature review chapter provides a comprehensive overview of the existing research and publications related to the topic of the thesis. It synthesizes and evaluates the current state of knowledge on the subject, identifies gaps in the literature, and provides a foundation for the research that follows.
- **Chapter 3** : In methodology chapter an in-depth discussion of overall research design, factors that can obstruct the research's developing phase, procedures of developing a model including data collection, data analysis, and classification methods used in the research, and research design or the proposed model is elaborated.
- **Chapter 4** : The evaluation and discussions chapter clearly describes the evaluation methodology, including dataset used, experimental setup and the simulation metrics employed(accuracy, sensitivity, specificity, precision and f1 score). The results including quantitative data, and any visualizations using tables and graphs are discussed. Also provided a comparitive study with other approaches to highlight the

quality of proposed model.

- **Chapter 5** : The conclusion summarizes the key findings from the evaluation and highlights the main contribution of the research. It provides an explanation for the research conduction, discussion on the model's performance over other existing models, the objectives achieved by the research and also discuss what could be the future works that can be conducted relating to the research.

2 Literature Review

2.1 Databases

- Google Scholar
- IEEE Explore
- Science Direct (Elsevier)

The above listed are well-known databases that is used in the study to get relevant related papers, and obtain information to examine trends in the field of study. Most recent literatures are studied and compared on their accuracy in sorting the real and fake reviews.

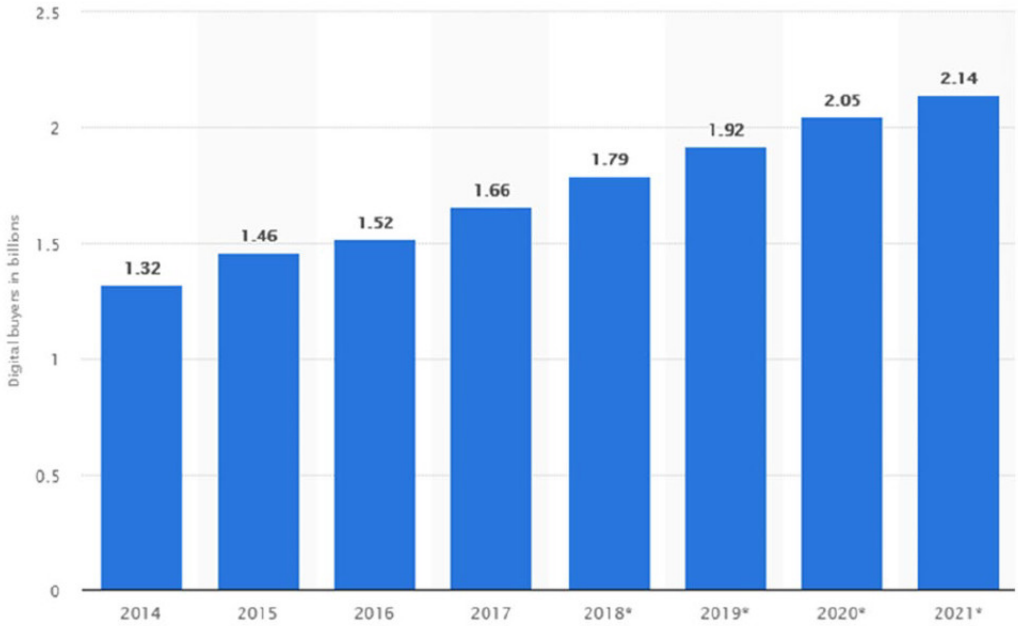


Figure 2.1: Worldwide digital buyers statistic in billions

2.2 Related Works on Review Analysis

Online reviews have become an integral part of the decision-making process of consumers, with a significant impact on businesses. Every day, many new businesses emerge from niche markets. According to the E-commerce Foundation, global company turnover reached $2671 billion in 2016, which was regarded a significant achievement in the sphere of online marketplace. According to reports, the proportion of digital buyers is steadily increasing. Statista published figures for digital buyer e-commerce in billions from 2014 to 2021, as well as retailing share from 2015 to 2021, as shown in figure 2.1 and 2.2. According to the survey, digital buyers will reach 2.14 billion by 2021, and the e-commerce retailing sector will reach 17.5% by 2021, which is more than double what it was in 2015, Altab et al. (2022).

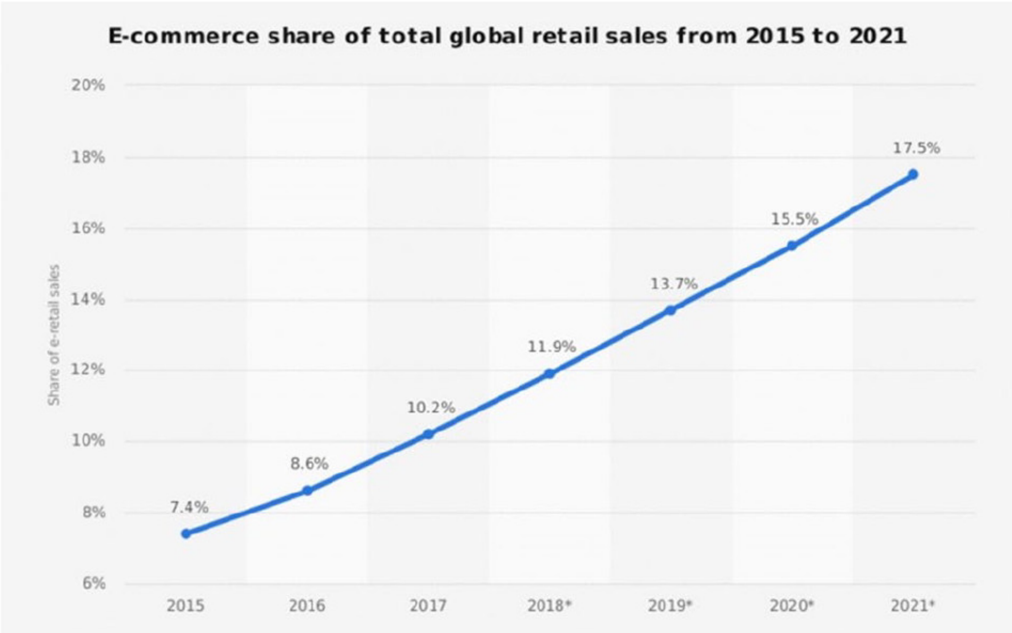


Figure 2.2: E-commerce share of total global retail sales from 2015 to 2021, Altab et al. (2022)

However, the proliferation of fake reviews has raised concerns regarding the credibility of online reviews. Detecting fake reviews has become a critical issue for e-commerce sites, businesses, and consumers. Several studies have been carried out recently to create models and algorithms to identify fake reviews on e-commerce websites.

Sihombing and Fong (2019) discussed that Yelp, a crowd-sourced review site frequently used by people to write reviews about the companies they have dealt with. The study found that there had been over 177 million reviews on the Yelp website by the end of 2018. For identifying fake reviews in Yelp dataset, the author performed Logistic Regression, Gaussian Naïve Bayes, Support Vector Machine, and XGBoost machine learning classification techniques. The experiment's result concluded that using XGBoost, the average F1 score for prediction was 99%, the highest. However the dataset's imbalance needs to be addressed.

Another study by Reddy et al. (2022), developed a classification system that can recognize what a customer posts on a hotel's website and feels about the hotel by applying sentiment analysis and machine learning algorithms to existing hotel reviews. As the classification algorithms, logistic regression and support vector machines (SVMs) were selected. The accuracy value for the logistic regression classifier model was 89.77%. The success rating for the Support Vector Machine classifier model was 86.49%. The classification model using logistic regression works better than the SVM model. It demonstrates how these techniques can be used to accurately categorise hotel reviews based on sentiment or opinion. The findings add to a better understanding of sentiment analysis in the context of hotel reviews and can help businesses in the hospitality industry improve customer experience and make decisions based on consumer input.

In ”Time is Important in Fake News Detection: a short review”, Rastogi and Bansal (2021) discussed on the time in the news cycle as a crucial factor. This study offers an early, middle, and late ”three-tier method” for identifying fake information. A comparison of the three tiers in terms of the goals, strategies, datasets, and characteristics discussed in the literature, see Figure 2.3. The outcome demonstrates how the desire to recognise fake news and the information or features that are readily accessible change over the course of a news cycle.

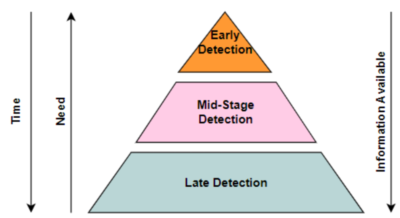


Figure 2.3: Three-tier fake news detection system, (Rastogi and Bansal (2021))

G et al. (2022) examined the effectiveness of various Deep Learning(DL) and Machine Learning(ML) model-based Fake News Detection(FND) approaches as well as the characteristics of various fake news contents. The analysis demonstrates that Natural

Language Processing(NLP) with DL models outperforms ML model-based identification. Additionally, the research found that the Bi-LSTM is the best DL model for FND. The study examines Naive Bayes, Support Vector Machines (SVM), Random Forest, and Neural Networks as machine learning techniques used for fake news detection. The writers outline the advantages and disadvantages of each strategy and provide insights into how they work in various settings. Furthermore, the research highlights the significance of feature selection and feature engineering in the identification of bogus news. It emphasises the importance of taking into account linguistic, semantic, and contextual factors in order to increase the accuracy and reliability of detection models.

”Fake Reviews Detection with Hybrid Features Using Time-Sequential Deep Learning Model” paper studied about using hybrid features, such as textual and behavioral ones, time-sequential neural network models can detect fake reviews, achieving accuracy levels that are close to the state-of-the-art. The most complete feature set, which contains 133 features, was developed by Satia Budhi et al. (2021), and the performance was exceptional, earning an F1 score of 84%. Weng et al. (2022) compared the effects of textual and behavioral characteristics, the F1 score was 89% and the ultimate accuracy score increased by 3%.

In order to identify fake evaluations using unsupervised learning, three models were used in this study by Mothukuri et al. (2022) to determine the best cluster. Gaussian mixture model(GMM) full covariance algorithm gave an accuracy of 74.6%, GMM diagonal covariance algorithm gave an accuracy of 64.4% and the k-means clustering model which separates the fake and non-fake reviews very clearly using the features giving an accuracy of 85% when tested with that of original results. Precisions of 0.86, 0.48, and 0.35 respectively had been observed in the k-means clustering model, GMM full covariance, and GMM diagonal covariance. Another article by Hassan and Islam (2019) presents semi-supervised and supervised classification methods for identifying fake online reviews. We employ the Expectation-maximization method was employed for semi-supervised learning. In our study, we use Support Vector Machines (SVM) and Statistical Naive Bayes classifiers as classifiers to enhance classification performance. By using supervised classification with a Naive Bayes classifier, we were able to boost the accuracy of semi-supervised classification to 85.21% and also discover the maximum accuracy of 86.32%.

In the research by Joshi and Abdelfattah (2021), six distinct machine learning models were put to the test to see how well they classified diseases based on drug reviews. The algorithms employed are listed below: Linear Support Vector Classifier (SVC), Multinomial Naive Bayes, Multinomial Logistic Regression, Decision Trees, Extra Trees, and Random Forests. According to its Precision, Recall, and F1 score, Linear SVC was discovered to be the most successful in forecasting the medical conditions. With the exception of Multinomial NB, Linear SVC's training and forecast times were all noticeably faster than those of the other classifiers. However, only the ten most prevalent conditions mentioned in the reviews were used to make these forecasts.

Guan (2022) analyzed the sentiment tendency of financial online reviews into negative reviews, positive reviews and neutral reviews, analyzed the application of Support Vector Machine and Naive Bayes classifier in review sentiment classification, and compared the performance of these two machine learning algorithms. In the experiment, it was discovered that the Naive Bayes classifier had a better classification accuracy, but machine learning algorithms are not very good at classifying neutral reviews. To summarize, this study makes a valuable contribution to the field by investigating the sentiment classification of financial online reviews through the utilization of machine learning algorithms. The results indicate that these algorithms possess the capability to accurately classify the sentiment conveyed in such reviews. The knowledge gained from this research holds significance for financial businesses as it enables them to derive meaningful insights from customer feedback and make well-informed decisions by employing sentiment analysis of online reviews.

Roshan R. Karwa (2022) presents a hybrid deep neural network model that combines Deep CNN and C-DSSM models, as in figure 2.4. It utilizes the LIAR dataset to recognize and categorize fake news. The proposed model achieved accuracy, recall, precision, and F1 score of 92.60%, 92.40%, and 92.50%, respectively, according to experimental findings. The performance of the proposed model is remarkable when it is compared to previous studies for identifying fake news using the LIAR dataset. As a result, the suggested hybrid model performs better when it comes to identifying and categorizing fake news on social media.

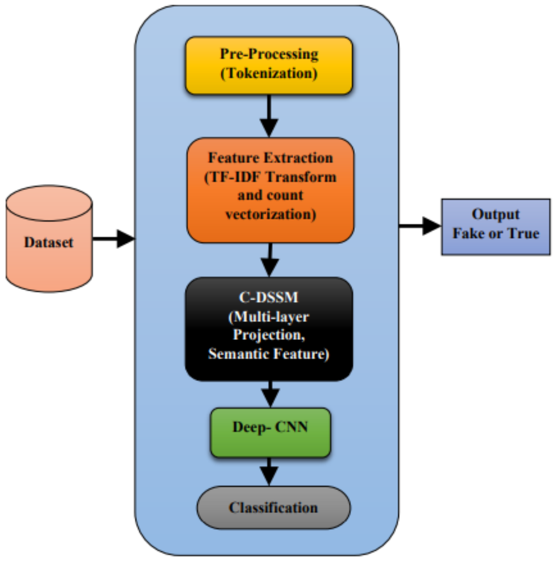


Figure 2.4: Automated Hybrid DCNN Model Architecture, (Roshan R. Karwa (2022))

Han and Mehta (2019), talks about fake news issues that seem to get worsen over time which leads to inaccurate perceptions of some information. The two types of news content—linguistic (including texts, headings, etc.) and visual—allow for differentiation between them. (including video and image-based). The most recent deep learning approaches, such as hybrid CNN and RNN, are evaluated to compare the performance with the traditional machine learning approaches. The traditional machine learning approaches, such as clustering and Naive Bayes, are evaluated for recognition accuracy and concluded the hybrid CNN and RNN outperforming the machine learning methods.

Traditional methods for evaluating spam views have drawbacks, such as lengthy computation times, insufficient accuracy, and an inability to effectively handle sparse data. Due to its self-adaptive nature, deep learning can effectively detect fake views. Deep learning techniques are known to work better than conventional machine learning algorithms is discussed by Bathla and Kumar (2021), in Opinion Spam Detection using Deep Learning. GloVe word encoding is used by CNN. The CNN model incorporates crucial text characteristics to increase accuracy. Furthermore, the Adam optimizer and dropout are used to integrate improvements.

In the study by R et al. (2022), Vader was used for sentiment analysis, the Gensim library was used for summarization, and the author used supervised machine learning method with SVD dimensionality reduction for the fake review classifier. The novel ideas from the study include using SVD dimensionality reduction and Logistic Regression Classifier to detect fake Amazon reviews and summarizing non-fake reviews after grouping them according to sentiment, achieving a performance rate of 81%. It emphasises the need of automating these chores, which could benefit industries that rely on online evaluations and comments.

Anas and Kumari (2021) proposes a system for detecting and removing fake reviews by analyzing the sentiment of reviews using opinion mining techniques. The proposed system uses Naive Bayes and Random Forest algorithm to classify reviews as positive, negative, or neutral, and then identifies reviews that are likely to be fake based on the sentiment and other features of the reviews. By a wide margin, the Random Forests model outperformed the Naive Bayes method by 89%.

To differentiate between spam opinions and authentic reviews, traditional machine learning methods like SVM, Logistic Regression, and Naive Bayes are used, according to Bathla and Kumar (2021) their accuracy is insufficient. Reviews text is observed and fed into a convolutional neural network (CNN) model in this study and to increase accuracy, the CNN model incorporates crucial text characteristics and GloVe word embedding is used. The proposed model outperformed traditional approaches with 92.81

2.3 Related Works on CNN-BiLSTM

According to Hu et al. (2019), CNN successfully extracts the inherent properties of historical urban water usage data, weather data, and meteorological data. The extracted data is sent into Bi-LSTM. Furthermore, Bi-LSTM fully reflects the long-term historical process as well as the future direction. Using the Bi-LSTM layer, you can avoid manually extracting a large number of features, which is common in standard machine learning methods. The CNN-Bi-LSTM model combines the benefits of the CNN and Bi-LSTM models, as well as the Bi-LSTM's nonlinear fitting capabilities to the time series model. When compared to these single models, the prediction accuracy of urban water forecast was enhanced. The accuracy of prediction of the CNN-Bi-LSTM model is 1% to 2% greater than that of the LSTM model, and it improved by 2% to 3% when compared to the CNN model. The prediction accuracy was improved by 0.6% to 1.5% when compared to the Bi-LSTM model.

As proposed by Ni et al. (2022), the CNN-Bi-LSTM model, accepts sequential ISAR

Table 2.1: Comparitive study on related works

S. No	Author and Year of Publication	Title of the article	Adopted methods	Authors Contribution	Pros and Cons	Accuracy
1	Weng et al., 2022	Fake Reviews Detection with Hybrid Features Using Time-Sequential Deep Learning Model	Long-Short Term Memory Deep Learning Model	Used textual and behavioral features in time-sequential neural network models to detect fake reviews	LSTM DL wasn't time consuming, and behavioral and textual feature of review combine to outperform the benchmark by 3. Levels and types of transparency-related explanations were not provided	89.62%
2	Mothukuri et al.,2022	Fake Review Detection using Unsupervised Learning	K-means clustering model, GMM full covariance algorithm and GMM diagonal covariance algorithm	Fake reviews identification using unsupervised learning. Accuracy is calculated using the original output given by yelp website.	Different boxplots were plotted by K-means grouping for better understanding of feature's behavior. Number of misclassifications were found low. More unsupervised learning methods need to be studied.	85%
3	Reddy et al., 2022	Classification of Hotel Reviews using Machine Learning Techniques	Logistic Regression and Support Vector Machine	Perform sentiment analysis for hotel reviews. Comparison of Logistic Regression and Support Vector Machine on given dataset is performed	Attempt to improve accuracy of the study need to be done. Data imbalances need to be addressed and more models need to be studied.	89.77%
4	Joshi and Abdelfattah, 2021	Multi-Class Text Classification Using Machine Learning Models for Online Drug Reviews	Linear Support Vector Classifier (SVC), Multinomial Naive Bayes, Multinomial Logistic Regression, Decision Trees, Extra Trees, and Random Forests	Find the most efficient model to predict the medical condition based on the users' reviews. Among 6 models Linear SVC outperformed.	Predictions were made only on a limited set of medical conditions, hence not deployable in real world. Study attempts to open that online reviews can predict or classify different forms of illness	88%
5	Sihombing and Fong, 2019	Fake Review Detection on Yelp Dataset Using Classification Techniques in Machine Learning	Logistic Regression Gaussian Naïve Bayes, Support Vector Machine, and XGBoost	Reviewed four popular machine learning classification methods for finding fake Yelp reviews analyzing the rating, the length of the reviews, their similarity, the reviewer's extreme rating ratio, the reviews' deviation from the mean rating, and the reviewer's maximum number of reviews per day	The dataset's limitations, such as the user confidence factor based on user friendship and user profile, prevent the implementation of additional features. Data imbalances need to handle.	99%
6	Han and Mehta, 2019	Fake News Detection in Social Networks Using Machine Learning and Deep Learning: Performance Evaluation	Naïve Bayes, Hybrid CNN and RNN	The performance of deep learning approaches, like hybrid CNN and RNN, is assessed in comparison to conventional machine learning approaches, clustering, Naive Bayes, for detection accuracy.	There is a need for research on choosing a machine learning or deep learning approach for issue solving that strikes a balance between accuracy and portability.	80%
7	Bathla and Kumar, 2021	Opinion Spam Detection using Deep Learning	Naïve Bayes, SVM, Logistic Regression, CNN with GloVe	Comparative analysis of the proposed method CNN with GloVe with traditional approaches	Reviewer behavior and other factors such as location, sentiments, entropy and ratings deviations need to be included in proposed model to enhance the accuracy.	92%
8	R et al., 2022	Detection and Summarization of Honest Reviews Using Text Mining	Stochastic Gradient Descent Classifier, Logistic Regression, SVM, Naïve Bayes	Summarize reviews after removing fake reviews to give the user a reliable assessment of the product.	To improve the categorization of fake reviews, features like review helpfulness can be considered.	81%
9	Anas and Kumari, 2021	Opinion Mining based Fake Product Review Monitoring and Removal System	Naïve Bayes and Random Forest	Detect and remove fake reviews by analyzing the sentiment of reviews using opinion mining techniques.	An algorithm may not successfully identify patterns or make reliable predictions about the reviews if it is not well-suited to the particular context of the reviews being analyzed.	89%
10	Bathla and Kumar, 2021	Opinion Spam Detection with Deep Learning	Convolutional Neural Network	The vector representation of words in traditional methods were replaced with text features in CNN model.	Replaced vector representation of words with essential features, improving accuracy. To enhance accuracy, reviewers location, behavior etc, could have been studied.	92%

images as input and outputs classification results, is trainable from start to finish. During the model learning and training process, feature extraction and target classification are carried out concurrently, which helps to extract deep key characteristics and enhance classification accuracy. The CNN-Bi-LSTM provides good classification accuracy, as shown in table 2.2 for distorted ISAR images and may be used to efficiently classify non-cooperative targets such as satellites and aircraft.

Table 2.2: Analysis by Ni et al. (2022)

Methods	Sequence length	Acc (%)
CNN	S=1	84.77
CNN-LSTM	S=3	88.72
	S=5	89.62
	S=10	90.83
CNN-Bi-LSTM	S=3	90.34
	S=5	91.73
	S=10	92.92

In contrast to the typical machine learning-based action classification approach, in this research Wu and Tang (2021) conducted study and offers a parallel CNN-BiLSTM neural network for user action feature extraction and classification. The experimental results show that the network can recognise user actions at a rate of 98.7%. When compared to support vector machines, signal CNNs, and signal BiLSTM networks, the parallel CNN-BiLSTM model classification technique offers high accuracy, great learning ability, and robustness even under diverse users and multiple action categories.

Table 2.3: Analysis by Wu and Tang (2021)

Methods	*SVM*	*CNN*	*BiLSTM*	*CNN-BiLSTM*
Accuracy	87%	90.7%	89.6%	98.7%

By developing a branched deep neural network that employs a shared feature space for both SAs and CAs, Tahvilian et al. (2022) studied two variations of a multi-task network, CNN-LSTM and CNN-BiLSTM, with shared parameters to recognise simple activities (SAs) and complicated activities (CAs) simultaneously. The study used 65 activities to train and evaluate the models: 51 SAs and 14 CAs made up of Lee Silverman Voice Treatment-BIG and functional activities and included 43 healthy volunteers, seven women and 36 males. The data were collected using four smart bands with embedded IMUs that were worn on both wrists and thighs. Our results reveal that the CNN-BiLSTM model beats the CNN-LSTM model, with average accuracies of 84.17% and 78.78% for SAs and CAs, respectively.

Table 2.4: Analysis on performance of CNN-BiLSTM

SL No.	Author and Year of Publication	Title of Article	Contribution
1	Hu et al. (2019)	A hybrid model based on CNN and Bi-LSTM for urban water demand prediction	CNN extract the inherent characteristics and Bi-LSTM fully reflect the long-term historical process and future trend of the data improving the performace by 1-2%
2	Ni et al.(2022)	Sequential ISAR Images Classification Using CNN-Bi-LSTM Method	CNN network extract the shallow features of ISAR images and Bi-LSTM network realizes the deep feature extraction and bidirectional feature fusion of the sequential ISAR images improving accuarcy by 2%
3	Wu and Tang(2021)	Research on User Action Recognition Method Based on parallel CNN-BiLSTM neural network	When compared to other networks, the parallel CNN-BiLSTM model classification technique offers high accuracy of 98.7% under diverse users and multiple action categories.
4	Tahvilian et al.(2022	Accuracy improvement in simple and complex Human Activity Recognition using a CNN-BiLSTM multi-task deep neural network	Compared CNN-LSTM and CNN-BiLSTM to recognize simple human activities (SAs) and complex human activities (CAs)

2.4 Research Gaps

In e-commerce platforms, machine learning (ML) or deep learning (DL) models are frequently used to detect false reviews. Based on characteristics including review length, sentiment, and rating, ML models like decision trees, random forests, and support vector machines (SVMs) and based on the textual content of the reviews, DL models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have also been used to identify false reviews.

A model efficient at extracting the long-term correlations in the text and retrieving local features from the text is important for detecting fake reviews. To better discriminate between authentic and false reviews, a model need to extract useful information from the reviews and also need to adapt to differences in the language and writing style of reviews.

The proposed study will evaluate CNN-BiLSTM's performance against other DL models like CNN and LSTM as well as more established ML models like Naive Bayes, and Random Forests. The research will make use of a dataset of reviews from e-commerce sites to evaluate each model's ability to identify fraudulent reviews in terms of accuracy, precision, recall, and F1-score.

2.5 Chapter Summary

Above discussed perspectives of researchers and their suggestions on different ML and DL models focus on the fact that no model stands alone with a numerically high standard accuracy in detecting the fake reviews in e-commerce platform. Failing in adapt to differences in the language and writing style of reviews, unable to identify the dependencies in the reviews, locally and globally, are some factors for the search of a better algorithm to detect fake reviews.

Reddy et al. (2022), Joshi and Abdelfattah (2021) and Sihombing and Fong (2019) discussed on different models and also came on with approximately 90% accuracy in prediction, but failed to handle data imbalance, data generalisation and the data-set limitation. Others by Weng et al. (2022), Bathla and Kumar (2021), Han and Mehta (2019) worked on the textual and behavioural nature of review, vectorisation of words and portability improving accuracy.

The studies points out that hybrid models have set good accuracy score in the detection. Analysing the existing methods and identifying the ways to improve the accuracy, CNN-BiLSTM, a hybrid from Convolutional Neural Network and Recurring Neural Network can perform in a better way considering the correlation and features in the reviews. Wu and Tang (2021), Ni et al. (2022), Hu et al. (2019) and Tahvilian et al. (2022) conducted research comparing CNN-BiLSTM with other models. As analysed in table 2.4, in each study performed CNN-BiLSTM outperformed other models including its base models CNN and LSTM as well.

3 Methodology

Through internet sites, applications, and services offered by companies, users provide reviews of electronics, cosmetics, and other products and services. We must read hundreds of reviews in order to analyse them, which requires a lot of time and analytical teams. It might be difficult and risky to identify the primary and recurring problems. However, there is little room for development as a result of these reviews. And here comes the role of a fake review detector, which identifies the real and fake reviews and make its much easier for the analytical team to analyse the trend or faults to be rectified and also makes it effortless for the users buying and selling products online to establish their brand authentically (Jain et al. (2022)).

While developing a model, it is necessary to analyse several factors to ensure effective and accurate system. Some factors include

- **Effect of availability of data** : A sufficient volume of labelled data that includes both real and fake reviews is crucial while considering the quality of training and evaluating a model.
- **Effect of model selection** : Analyse how model complexity, interpretability, and performance trade off. Simple models like Naive Bayes or Logistic Regression might not be able to analyse and capture the complex pattern that models with deep learning architecture or ensemble models can.
- **Evaluation Metrics** : Establish suitable evaluation measures to evaluate the effectiveness of the model. Utilise appropriate metrics to evaluate the models while taking your application's unique requirements into account. Consider measurements like recall or F1 score if you wish to prioritise recall (reduce false negatives).
- **Generalisation** : Examine how well the model perform on new data. Assess the model's performance using different test datasets or cross-validation methods to determine how well it can detect fake reviews in new samples.

The proposed approach of implementing models workflow has four phases, as shown in figure 3.1 :

1. Data Exploration
2. Data Manipulation
3. Data Pre-processing
4. Classification of the reviews

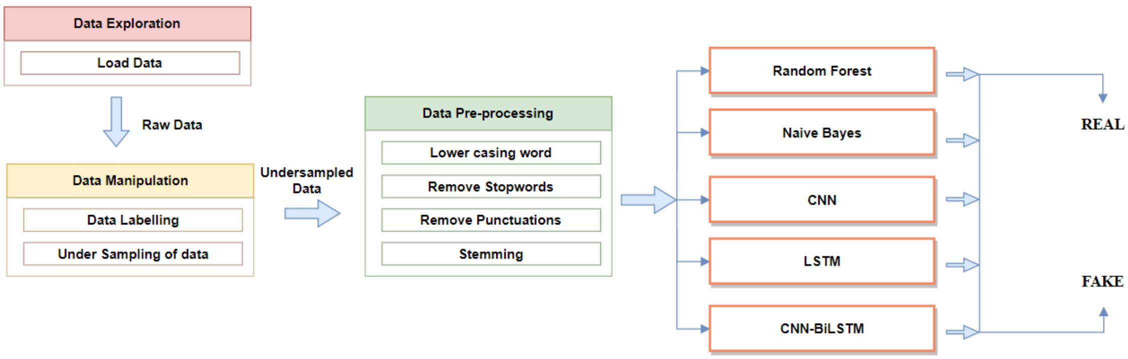


Figure 3.1: Proposed Workflow Diagram

3.1 Data Exploration and Manipulation

The selection of dataset for the system is crucial for effective implementation of a model. This involves examining the distribution of the classes, conducting exploratory data analysis, and performing validation experiments to ensure the dataset is suitable for your specific task of fake product review detection. As discussed by N and Gupta (2020), there are many factors that can affect the selection of dataset, and few are:

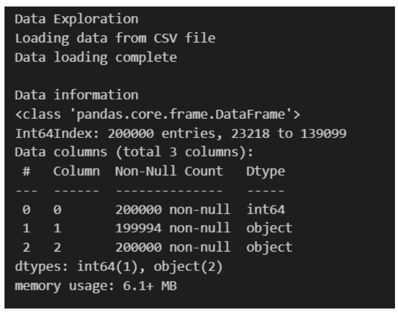


Figure 3.2: Data Information

- Size and diversity : The dataset chosen should be large and diverse, which can help the model learn widely and ensure generalised model training.

- Labelling and Ground Truths : The dataset should be reliable and accurately labelled against the review as real or fake.
- Quality of Reviews: Pay attention to the quality of reviews in the dataset. Ensure that the reviews are coherent, well-written, and representative of real-world scenarios.
- Data Source: Consider the source of the dataset and its relevance to the target domain. Ideally the dataset should be collected from e-commerce platforms or reputable sources known for their authentic product reviews.

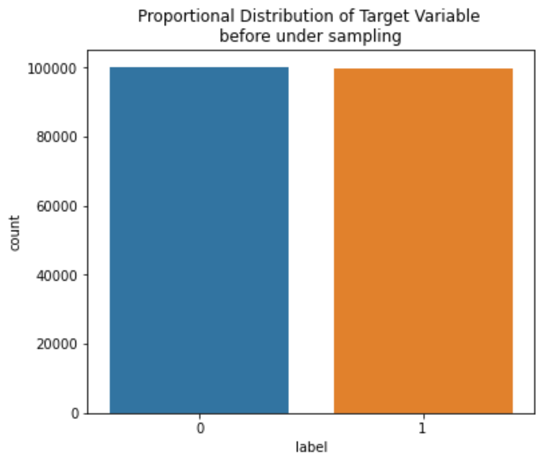


Figure 3.3: Proportional Distribution of Target Variable before sampling

To begin, data need to be loaded and then studied. The research uses amazon reviews from Kaggle dataset. The data is initially read from test.csv file with 4 lakh reviews, and 50% of the reviews are considered for the study. The dataset has polarity of the reviews mentioned using which the labelling of data is done. Also the review title and content are two separate columns, which is then combined. A crucial problem that exists for the study is data imbalance and use of sampling techniques improves the efficiency of the classifier (Basha et al. (2022)).

The proportional distribution of the target variable before and after sampling of the dataset are as shown in figure 3.3, 3.4. Since the count of real and fake reviews equals the system didn't perform under sampling of real reviews.

```
Initiating undersampling of data
 Real review count: 99823
 Fake review count: 100177
Require no under sampling
```

Figure 3.4: Under Sampling of Data

3.2 Data Preprocessing

Data preprocessing is an important step in preparing data for ML or DL models. To make sure the raw data is of high quality, consistent, and appropriate for the selected model, it must be transformed and cleaned. It also helps in mitigating data-related issues, preparing the data for specific analytical techniques, and enhancing the performance and interpretability of the resulting models or insights. The preprocessing stage has the following steps to make the text predictable and analyzable (Hafeez and Kathirisetty (2022)).

- **Case Folding** : All the texts in the review are cnverted to lowercase in order to reduce the dimensionality.
- **Removal of punctuation and stopwords** : To eliminate noise, all punctuation and stopwords from the dataset were eliminated and replaced with spaces.
- **Tokenization** : Tokenization breaks the words and identify the root key or phrases.
- **Stemming** : Stemming is done on the remaining words to only extract the root words from the corpus as features

Vector Generation : Term Frequency-Inverse Document Frequency(TF-IDF) and Keras text tokenizer are tools used in natural language processing (NLP) tasks for the study.

- **TF-IDF** : TF evaluates how frequently a term appears in a a text, while IDF evaluates a terms rarity over the entire corpus. It provides a numerical representation of text documents, enabling machine learning algorithms to work with text data and also it identifies most relevant and distinct features for text classification (Bakiyev (2022)). A term's TF-IDF value in a document is calculated by multiplying its TF by its IDF. This computation highlights phrases that are both common within a document and relatively uncommon across the entire corpus. High TF-IDF scores suggest phrases that are significant and unique to a given document.
- **Keras text tokenizer** : An effective method for transforming text into numerical sequences that may be input into a neural network is the Keras text tokenizer. It is frequently applied to NLP tasks in deep learning models (Kapali et al. (2022)).It converts text into numerical representations, making deep learning models easier

to train and use for applications like sentiment analysis, text categorization, machine translation, and more.

3.3 Classification Phase

When product reviews are properly categorised, they can be used to identify fake ones. This section discusses Random Forest, Naive Bayes, CNN, LSTM and CNN-BiLSTM, to categorise the fake and real reviews.

1. **Random Forest**

   The Random Forest algorithm is highly effective for review detection tasks, offering a range of benefits including robustness, feature importance analysis, strong performance, and interpretability. As a supervised learning approach, it is used to develop and evaluate machine learning models and belongs to the ensemble learning family. In order to enhance performance and improve prediction accuracy, Random Forest generates multiple decision trees and combines their predictions (Dutta et al. (2020)).

   To apply Random Forest in the context of review detection, the reviews undergo an initial vectorization process and are then split into a training set and a test set using the TF-IDF (Term Frequency-Inverse Document Frequency) technique. The implementation of the Random Forest Classifier provided by Scikit-learn, a popular machine learning library, is utilized for the classification of the reviews. The model is trained using the training data and subsequently evaluated using the test data. The predicted outputs are compared to the actual labels to validate and assess the performance of the model.

2. **Naive Bayes**

   Naive Bayes may appear to be a basic classification algorithm, it plays a crucial role in predictive modeling due to its simplicity and effectiveness. It belongs to the family of probabilistic classifiers and relies on conditional probabilities. Various versions of Naive Bayes, such as Gaussian, Multinomial, and Bernoulli, exist. Multinomial Naive Bayes is commonly employed in text categorization tasks, particularly when dealing with discrete features like word counts or frequencies (Bhatia and Malhotra (2021)).

Multinomial Naive Bayes is specifically designed to handle discrete features extracted from preprocessed review data, such as word frequencies or counts. The algorithm assumes that these features are independent of each other given the class label. During the training phase, labeled review data is utilized. The algorithm estimates the probabilities of each class label (e.g., positive or negative) and the conditional probabilities of each feature (e.g., word) given the class label. This estimation is accomplished by calculating the relative frequencies of features within each class. Once the model is trained using the labeled data, it can be applied to predict the class labels of new and unseen reviews. By considering the probabilities of different class labels and the conditional probabilities of features, the algorithm assigns the most probable class label to each review. This process facilitates the classification of reviews into categories such as positive or negative sentiment.

Naive Bayes algorithms, including Multinomial Naive Bayes, are popular choices for text classification tasks due to their simplicity, computational efficiency, and ability to effectively handle discrete features. Despite the ”naive” assumption of feature independence, Naive Bayes classifiers often achieve impressive results in practice. They find widespread use in various applications, including spam filtering, sentiment analysis, and document classification, where discrete features are prevalent.

3. **CNN**

   Compared to standard neural networks, CNN models have a unique architecture that enables them to effectively identify spatial and local correlations in textual data, making them well-suited for false review detection. Popular frameworks like TensorFlow and Keras provide user-friendly APIs that simplify the creation and training of CNN models specifically designed for tasks such as text categorization.

   In the context of false review detection, the CNN model takes a list of word embeddings as input (Dai (2021)), which represent words as dense numerical vectors capturing their semantic meaning. Through the use of convolutional layers with various filter sizes, the model identifies regional patterns and n-gram features within the text. These convolutional layers produce feature maps that are then downsampled using techniques like max pooling or global pooling. For improved performance, additional layers such as dropout, batch normalization, or recurrent layers can be selectively added to the model. These layers introduce regularization, speed up training, and enhance the model’s ability to generalize to new data. During the training process, the model learns to recognize significant features in the input text and make predictions based on those features. To train the CNN model effectively, suitable optimization methods like Adam or stochastic gradient descent (SGD) are used. These methods iteratively adjust the model's weights to minimize the loss function, typically measured using categorical cross-entropy. By optimizing the model with appropriate techniques and loss functions, the CNN becomes proficient at accurately predicting outcomes based on the input text.

In conclusion, a CNN model designed for false review detection leverages its specialized architecture and the ability to identify spatial and local correlations in textual data. Through training, the model acquires the capability to identify important features and make accurate predictions. Careful selection of layers, optimization methods(like Adam or SGD), and loss functions (such categor- ical cross-entropy) significantly enhance the CNN model's performance in detecting false reviews(Wang and Gang (2018)).

4. **LSTM**

According to Shoryu et al. (2021), deep learning methods play a crucial role in the current approach to emotion-based text categorization, specifically through the utilization of LSTM (Long Short Term Memory), a type of recurrent neural network algorithm. LSTM models excel at capturing long-term dependencies in sequential data by employing memory cells for information storage and retrieval. Unlike traditional feedforward neural networks, LSTM networks possess a more complex structure with three key components: an input gate, a forget gate, and an output gate. These gates control the flow of information, allowing the network to retain relevant information and discard irrelevant details.

LSTM models have demonstrated notable success in various natural language processing (NLP) tasks, including sentiment analysis, text classification, named entity identification, and machine translation. Their ability to effectively model sequential data and capture contextual dependencies makes them well-suited for analyzing and comprehending textual information. The development and training of LSTM models for text classification have been made more accessible through the availability of practical APIs provided by frameworks such as TensorFlow and Keras. These APIs streamline the implementation process and facilitate the creation of LSTM-based models for analyzing textual data.

In summary, deep learning techniques, particularly LSTM models, are widely employed in the present approach to emotion-based text categorization. The distinctive ability of LSTM to capture long-term dependencies and leverage memory cells for information storage makes it highly effective in analyzing sequential data. Successful applications of LSTM models have been observed in various NLP tasks. The availability of practical APIs in frameworks like TensorFlow and Keras simplifies the development and training of LSTM models for text classification purposes.

3.4 The Proposed Model - CNN-BiLSTM

Implemetation and training of hybrid models require more processing power, specialised knowledge and careful integration of components. However they showcase the combined power of the constituent models.

- Captures complex patterns in data, resulting improved performance.
- Robust and Generalized model
- Handling complexity of data more effectively
- Provide flexibility in implemetation of a specific requirement

The above discussed are the advantages of using a hybrid model.
The literature review on the proposed model, as discussed in section 2.3 enhances the quality and efficiency of the chosen model CNN-BiLSTM. As mentioned by Hu et al. (2019), CNN successfully extracts the inherent properties from the data. The extracted data is sent into Bi-LSTM. And, Bi-LSTM fully reflects the long-term historical process as well as the future direction. Using the Bi-LSTM layer, you can avoid manually extracting a large number of features, which is common in standard machine learning methods. This fairly defines the prime reason for the selection of the model CNN-BiLSTM.

CNN is a feed-forward network which extracts useful details like local patterns from a vast set of data. To a window of size k, CNN applies a non-linear learning function. The window glides in order to complete each time step and reach the entire sequence. Therefore, it gives back a value that embodies the information from that time span. An n-dimensional vector is produced by applying a filter to each window. A pooling procedure is used to combine the resulting vectors into a single vector. A view of the entire sequence is represented by the vector. Unfortunately, CNNs are unable to handle an extensive global pattern of sequences (Behzadidoost and Izadkhah (2022)).

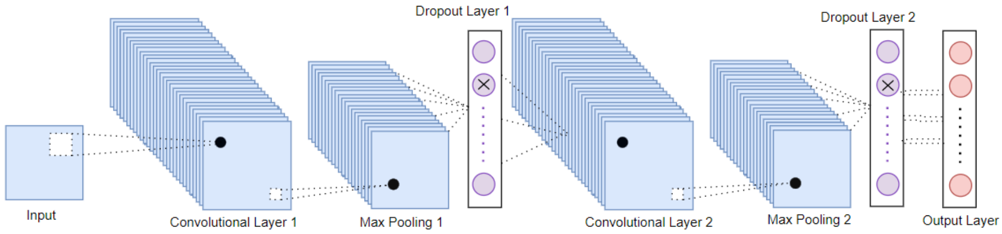


Figure 3.5: CNN Architecture

Recurring Neural Networks (RNN) are employed to analyse those dependencies among the tokens in the sequences. The previous inputs and states are stored in RNN using an internal gate. Recurrent neural networks can eliminate backpropagation error signals that flow in the opposite direction. The disappearing gradient is the name given to it. A remedy for the vanishing gradient is to use LSTMs. The input and forget gates of the LSTMs allow them to store the error flow constant. How much data can be saved is determined by these gates.A type of a neural network is a bidirectional LSTM (BiLSTM), which takes into account both the forward and backward directions, unlike LSTMs, which only take into account one direction during training (Behzadidoost and Izadkhah (2022)).

In the proposed model, the CNN architecture consists of six layers, 2 1D Convolution layer, 2 intermediate pooling layers and 2 dropout units, shown in figure 3.5. The use of 2 1D convolution layer adds to better feature extraction. Local patterns or edges are examples of lower-level characteristics that the first convolutional layer collects. The second convolutional layer relies on these features to capture global or long-range relationships by covering a broader receptive field. The model may take into account a wider context and gather more detailed information from the input sequence because to the extended receptive field (Zhang et al. (2018)).

As discussed by Behzadidoost and Izadkhah (2022), Correa-Delval et al. (2021) and other studies, LSTM helps in tackling the issue of vanishing long-term dependencies. The performance of the LSTM can be improved to further extends by bidirectional layers. Two parallel LSTMs make up bidirectional LSTM as shown in figure 3.6, one of which reads the input sequence forward and the other of which reads it backward. The output from the backwards and forwards parts is then combined for the final classification. BiLSTM

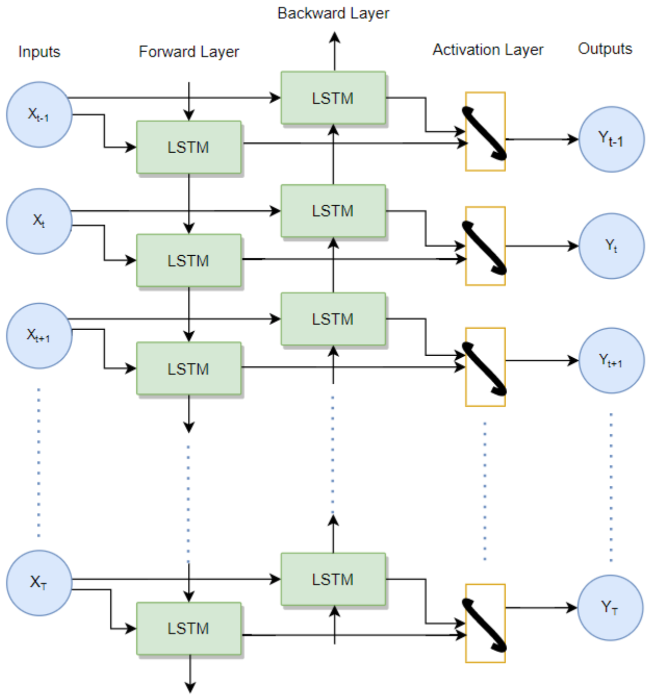


Figure 3.6: BiLSTM Architecture, Reddy and Ashwin Kumar (2022)

have a better comprehension of the input sequence by taking into account both the past and the future context, which leads to increased accuracy in text classification.

4 Evaluation and Discussions

Evaluation of a model's performance is an integral part of proposing a model. It is the way of convincing the world how accurate the model proposed is in solving the corresponding problem. This chapter focus on assessing the performance and effectiveness of the proposed model to accurately identify the fake reviews and distinguish them from the real ones.

The models studied in the research paper uses Amazon reviews from Kaggle dataset containing real and fake reviews of 400000 count. The evaluation is carried out on 2 lakh reviews. And in this research Confusion Matrix, Area Under the Receiver Operating Characteristic Curve(AUC-ROC) and Classification Report containing metrics like accuracy, precision, F1 score are used to analyse the performance of the proposed model. The training and evaluation of the proposed model and existing models using the dataset are performed, see Appendix E for the quantitative data, and visualizations using tables and graphs of the performance of each model studied in this research.

4.1 Parameters of Proposed Model

The parameters capture the underlying patterns and relationships in the data and are adjusted iteratively to minimize the model's loss or error. In the study, the parameters and specifications allowed to customize the architecture and behavior of the CNN-BiLSTM model for the detection of fake product reviews from the set of data in hand. Tuning these parameters can help improve the model's performance, accuracy, and ability to learn

Table 4.1: Parameter of CNN

CNN Parameters	
Convolutional Layers	2
Pooling Layers	Max Pooling
Activation Function	ReLU
Dropout Rate	0.2

Table 4.2: Parameter of BiLSTM

BiLSTM Parameters	
LSTM Layers	2
LSTM units	64
Activation Function	Linear
Dropout Rate	0.2

meaningful representations from the input data. Table 4.1 and 4.2 mentions the parameters or specifications for CNN and BiLSTM model's architecture, and table 4.3 provides the parameters or specifications of the overall hybrid model proposed which includes significant details about the model.

Table 4.3: Optimization Parameters of proposed model

Model	Parameters	Values
CNN - Bi LSTM	Network Configuration	Fully Connected
	Optimizer Algorithm	Adam
	Loss Function	Categorical Cross Entropy
	Total Epochs	10
	Rate of learning	0.01
	Activation	Softmax
	Fully connected layer	Dropout
	Batch Size	32
	Pool size	Max Pooling

4.2 Simulation Measures

Simulation measures or evaluation metrics are used to assess the performance and quality of a simulation model. These measures help in evaluating how well the simulation model represents the real system and how accurately it captures the desired behavior or outcomes. The measures to evaluate the model used in the study are Accuracy, Sensitivity, Specificity, Precision and F1 Score.

- Accuracy: The overall accuracy of the model, calculated as (TP + TN) / (TP + TN + FP + FN).
- Sensitivity (Recall or True Positive Rate): The proportion of true positive predictions among all actual positive instances, calculated as TP / (TP + FN).
- Specificity (True Negative Rate): The proportion of true negative predictions among all actual negative instances, calculated as TN / (TN + FP).
- Precision: The proportion of true positive predictions among all positive predictions, calculated as TP / (TP + FP).
- F1 Score: A metric that combines precision and recall, calculated as 2 * (Precision * Recall) / (Precision + Recall).

The mathematical expressions for calculating the about measures (equation 4.1 to 4.5) are discussed below.

$$Accuracy(Ac) = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

$$Sensitivity(Se) = \frac{TP}{TP + FN} \tag{4.2}$$

$$Specificity(Sp) = \frac{TN}{TN + FP} \tag{4.3}$$

$$Precision(P) = \frac{TP}{TP + FP} \tag{4.4}$$

$$F1Score(F1s) = \frac{2TP}{2TP + FP + FN} \tag{4.5}$$

Table 4.4 provides the confusion matrix for the proposed model, which has the details of True Positives(TP), True Negatives(TN), False Positives(FP) and False Negatives(FN).

Table 4.4: Confusion Matrix of Proposed Model

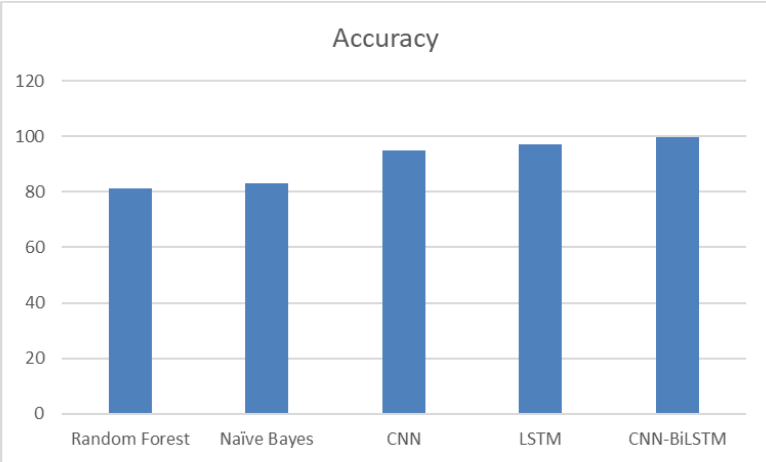
		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	20033	75
	Negative (0)	39	19862

4.3 Performance Analysis

The models studied in the research paper uses Amazon reviews containing real and fake reviews of 400000 count. 50% of data from the whole dataset was chosen to use in the research, ie, 200000 reviews. Among the chosen reviews each model used 80% data for training purpose of the models and 20% data for testing purpose. Among the 80% training data, in case of CNN, LSTM and CNN-BiLSTM, the taring dataset is further split into 80% for training and 20% for validation of the model, which follows the model evaluation by the test data.

The whole data set is labelled, and is mixture of real and fake reviews. The training sets labels are passed to the model for training itself and the validation set is internally split from the training data passsed by the model for self analysis on its accuracy and loss. The testing set is labelled and is kept unaware from the model. The labels of the

Figure 4.1: Model Comparison



test set reviews are then compared with the model's predicted labels. The test sets helps in the model evaluation and performance study.

Models	Accuracy
Random Forest	81.26
Naïve Bayes	82.96
CNN	95.09
LSTM	97.14
CNN-BiLSTM	99.73

Table 4.5: Comparitive Analysis

Comparitive study of the models, Random Forest, Naive Bayes, CNN, LSTM and the proposed CNN-BiLSTM is done with their accuracy achieved while evaluation of trained models are described in the table 4.5. Figure 4.1 illustrates the accuracies related to the models studied in the research. The acuuracy achieved by the Random Forest and Naive Bayes models are 81.26% and 82.96%, similar to the range of accuracy proposed models by Anas and Kumari (2021). The accuracy for CNN and LSTM are 95.09%, 97.14% respectively. Eventhough CNN and LSTM performed better, their hybrid CNN-BiLSTM, the proposed model gives an improved accuracy of 99.73%, which is quantitatively higher when compared with others. The quantitative data shown in table 4.5 clearly states the efficiency of the proposed model from the compared other models in the study. Refer appendix E, for detailed evaluation results.

The performance analysis of the proposed model for detecting fake product reviews are shown in fig 4.2. The figure visually represent the different evaluation metrics of

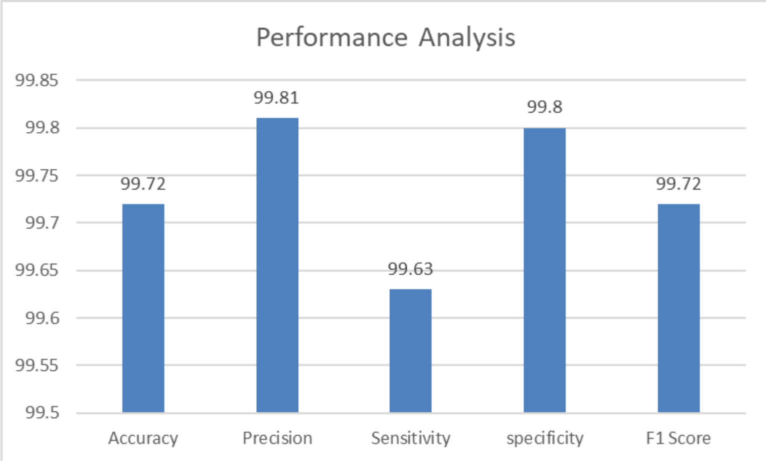


Figure 4.2: Performance Analysis

the proposed model, which provides a better understanding of the performance and efficiency of the model. The evaluation uses various simulation metrics, including accuracy, sensitivity, specificity, precision and F1 score. The research points out that the proposed model, CNN-BiLSTM model proves to obtain high rate of accuracy(99.72%), sensitivity(99.63%), specificity(99.80%), precision(99.81%) and F1 score(99.72%) in terms of the evaluation metrics, as depicted in fig 4.2.

The consistently high simulation measures indicates that the model is performing exceptionally well, achieving good balance between accurately identifying the true cases avoiding the false positives and false negatives in the classification. It also indicated the model's potential in generalizing the unseen data, and it's robustness and reliability as well.

5 Conclusion

Online reviews help consumers shop online more easily in the e-commerce sector, and as a result, customers are relying more on review data to assess the quality of goods and make purchasing decisions. The survey conducted by Altab et al. (2022), discusses on the growth of retail sales in e-commerce in year 2015 to 2021. By utilising the audience's genuine voice, encouraging consumers to post product reviews has evolved into a more effective technique for brand promotion in the modern world.

The study looks at different strategies for detecting fake product reviews including data preprocessing, feature extraction and classifications such as Random Forest, Naive Bayes, CNN, LSTM, and CNN-BiLSTM. Techniques used in data preprocessing includes case folding, removal of noise (punctuations, stopwords) and identifying root key or phrases. In next phase, combined title and reviews to form a single feature and labelled data with the polarity values available. Following performed analysis of the data using different classifiers, and evaluated their performance. The accuracy, sensitivity, specificity, precision and F1 score of models are considered to evaluate the overall performance of the proposed model.

5.1 Achievements on Aims and Objectives

- Developed a hybrid deep learning model of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), CNN-BiLSTM which can detect complicated textual patterns, helps to detect fake reviews with improve accuracy of 99.72%. The accuracy achieved by the proposed model is 4% higher than the CNN model alone and 2.5% higher than the LSTM model alone.

- Addressing class imbalance is necessary in machine learning tasks when the distribution of classes in the dataset is significantly skewed, meaning that one class has a much larger number of instances compared to the other class(es). This class imbalance can pose challenges and impact the performance of machine learning models in several ways such as bias towards the majority class, suppressing the minority class, fails to generalise the instances of the minority class data. These all leads to lower predictive performance of the model. This issue of addressing class imbalance in the data is eliminated using under sampling technique, that is by reducing the instances of majority class equal to the minority class instances during data manipulation for the analysis. The algorithm also checks the requirement for resampling.

- Evaluation of the performance of the proposed model is done on a publicly available dataset from kaggle. The training data and testing data are completely two different sets. The discussion is provided in detail in section 4 for more reference, and the results in appendix E. Figure 4.2 visualises the evaluation metrics used to evaluate the model.

- Compared the proposed model to other state-of-the-art machine or deep learning models. The proposed model by the research stands the better performed when compared with other models chosen. It is done and the comparison is shown in table 4.5.

The ultimate aim of the study is to improve the quality of customer reviews in various industries, and to achieve the same a system to monitor the fake products review is developed with improved efficiency when compared with other existing models as well its constituent models when stand alone.

5.2 Achievements on Research Questions

1. What is the proposed hybrid deep learning model techniques to detect fake reviews?

   The study developed a hybrid deep learning model with CNN and LSTM (RNN) which helps to detect fake reviews, by detecting complicated textual patterns and improved accuracy. CNN-BiLSTM outperforms its base models(CNN, LSTM) and other state-of-the-art like Random Forest and Naive Bayes, see table 4.5 for the comparitive study performed on the models.

2. How does the proposed model address the issue of class imbalance in fake review detection?

   Issue of class imbalance in the data is addressed by studying and applying under-sampling technique, achieved by resampling the real reviews to equal the fake review count.

3. How does the proposed model perform on a publicly available data-set in comparison to other state-of-the-art machine or deep learning models?

   The performance of the model proposed is initially validated using training data and evaluated using unseen data to generalize the models capability to identify fake reviews. Different metrics are used in evaluation such as accuracy, sensitivity, specificity, precision and F1 score.

4. What is the accuracy and efficiency of the proposed fake review detection system?

   Critically examined past researchs, including identifying and understanding limitations or shortcomings of previous studies, refer to table 2.1 for details. Developed a plan to address these issues which is more nuanced and well-informed, and design a study that is more likely to produce meaningful and impactful results. Different models such as Random Forest, Naive Bayes, CNN, LSTM and the proposed model CNN-BiLSTM are validated and evaluated, followed by a comparitive study of the models as shown in table 4.5 to analyse the relative performance of the models. The study evaluated the model with dataset from various domain which improved the generalisation of the model.

5. In what industries can the proposed system be used to improve the quality of customer reviews?

   Fake review detection can be used in various industries such as e-commerce, electronics, hospitality, food and beverage, and tourism to improve the quality of customer reviews. In the evolving world, online reviews contributes to the brand recommendation and promotion.

5.3 Future Works

The need for fake product review detection is fairly addressed in the study. The purpose to choose an algorithm to detect the fake product reviews is fulfilled. In future works, to further improve the performance and efficiency of the proposed model CNN-BiLSTM, the following suggestions can be considered.

One avenue to explore is the refinement of hyperparameter values in the CNN-BiLSTM model. This process, known as fine-tuning, entails systematically adjusting parameters like learning rate, batch size, choice of optimizer, and regularization techniques to optimize the model's performance. By carefully tuning these hyperparameters, the model has the potential to achieve higher accuracy and improved generalization.

Another area for improvement involves experimenting with diverse network designs. Modifying the architecture of the CNN-BiLSTM model by varying the number of LSTM or convolutional layer units can significantly impact its performance. Additionally, exploring different dropout rates and activation functions can effectively address overfitting issues and enhance the model's capacity to generalize well to new data.

Implementing a real-time fake product review detection system on production environment where the system automatically analyse and flag the fake reviews submitted. Real-time implementation requires efficient processing and scalability considerations to handle a large volume of reviews effectively.

References

Al-khiza'ay, M., Alallaq, N., Al-Mansoori, A. and Al-Sudani, A. R., 2019. Personalized reviews based on aspect analysis and polarity. *2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO)*, 1–6.

Altab, H. M., Yinping, M., Sajjad, H. M., Frimpong, A. N. K., Frempong, M. F. and Adu-Yeboah, S. S., 2022. Understanding online consumer textual reviews and rating: Review length with moderated multiple regression analysis approach. *SAGE Open*, 12 (2), 21582440221104806. URL `https://doi.org/10.1177/21582440221104806`.

AlZu'bi, S., Alsmadiv, A., AlQatawneh, S., Al-Ayyoub, M., Hawashin, B. and Jararweh, Y., 2019. A brief analysis of amazon online reviews. *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 555–560.

Anas, S. M. and Kumari, S., 2021. Opinion mining based fake product review monitoring and removal system. *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, 985–988.

Bakiyev, B., 2022. Method for determining the similarity of text documents for the kazakh language, taking into account synonyms: Extension to tf-idf. *2022 International Conference on Smart Information Systems and Technologies (SIST)*, 1–6.

Basha, S. J., Madala, S. R., Vivek, K., Kumar, E. S. and Ammannamma, T., 2022. A review on imbalanced data classification techniques. *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, 1–6.

Bathla, G. and Kumar, A., 2021. Opinion spam detection using deep learning. *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, 1160–1164.

Behzadidoost, R. and Izadkhah, H., 2022. An ensembled text-based multi-stacked-cnn-bilstm for rumor detection on social media. *2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, 1–6.

Bhatia, S. and Malhotra, J., 2021. Naïve bayes classifier for predicting the novel coronavirus. *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 880–883.

Chen, L. and Li, Q., 2022. How does fake review influence e-commerce platform revenue? *2022 23rd Asia-Pacific Network Operations and Management Symposium (APNOMS)*, 1–4.

Correa-Delval, M., Sun, H., Matthews, P. C. and Jiang, J., 2021. Appliance classification using bilstm neural networks and feature extraction. *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 1–5.

Dai, D., 2021. An introduction of cnn: Models and training on neural network models. *2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR)*, 135–138.

Dutta, K. K., S, S. A., Victor, A., Nathu, A. G., Ayman Habib, M. and Parashar, D., 2020. Kannada alphabets recognition using decision tree and random forest models. *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 534–541.

G, P., Germanaus Alex, M. and John Peter, S., 2022. Review of fake news detection in social media using machine learning techniques. *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, 496–501.

Guan, D., 2022. Sentiment classification of financial online reviews based on machine learning algorithm. *2022 International Conference on Education, Network and Information Technology (ICENIT)*, 289–293.

Hafeez, S. and Kathirisetty, N., 2022. Effects and comparison of different data preprocessing techniques and ml and deep learning models for sentiment analysis: Svm, knn, pca with svm and cnn. *2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR)*, 1–6.

Halim, E., Muzakir, Z. F. and Hebrard, M., 2022. The impact on review credibility and trust from review solicitation on e-commerce. *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 196–201.

Han, W. and Mehta, V., 2019. Fake news detection in social networks using machine learning and deep learning: Performance evaluation. *2019 IEEE International Conference on Industrial Internet (ICII)*, 375–380.

Hassan, R. and Islam, M. R., 2019. Detection of fake online reviews using semi-supervised and supervised learning. *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 1–5.

Hu, P., Tong, J., Wang, J., Yang, Y. and Oliveira Turci, L. d., 2019. A hybrid model based on cnn and bi-lstm for urban water demand prediction. *2019 IEEE Congress on Evolutionary Computation (CEC)*, 1088–1094.

Jain, P., Sharma, S., Monica and Aggarwal, P. K., 2022. Classifying fake news detection using svm, naive bayes and lstm. *2022 12th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, 460–464.

Joshi, S. and Abdelfattah, E., 2021. Multi-class text classification using machine learning models for online drug reviews. *2021 IEEE World AI IoT Congress (AIIoT)*, 0262–0267.

Kapali, N., Tuhin, T., Pramanik, A., Rahman, M. S. and Noori, S. R. H., 2022. Sentiment analysis of facebook and youtube bengali comments using lstm and bi-lstm. *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–6.

Mothukuri, R., Aasritha, A., Maremalla, K. C., Pokala, K. N. and Perumalla, G. K., 2022. Fake review detection using unsupervised learning. *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 119–125.

N, T. R. and Gupta, R., 2020. Feature selection techniques and its importance in machine learning: A survey. *2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS)*, 1–6.

Nagamma, P., Pruthvi, H. R., Nisha, K. K. and Shwetha, N. H., 2015. An improved sentiment analysis of online movie reviews based on clustering for box-office prediction. *International Conference on Computing, Communication Automation*, 933–937.

Ni, P., Sheng, J., Jiang, L. and Xu, G., 2022. Sequential isar images classification using cnn-bi-lstm method. *2022 3rd China International SAR Symposium (CISS)*, 1–5.

R, S. R., L, M. R., S, D., P K, A. D. and S, A. M., 2022. Detection and summarization of honest reviews using text mining. *2022 8th International Conference on Smart Structures and Systems (ICSSS)*, 01–05.

Rastogi, S. and Bansal, D., 2021. Time is important in fake news detection: a short review. *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, 1441–1443.

Reddy, R. and Ashwin Kumar, U. M., 2022. Convolutional neural networks- bidirectional long short term memory based fake review classification. *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 1069–1073.

Reddy, Y. C. A. P., Sagar, S. P. P., Kalyan, R. P. and Charan, N. S., 2022. Classification of hotel reviews using machine learning techniques. *2022 8th International Conference on Smart Structures and Systems (ICSSS)*, 1–5.

Roshan R. Karwa, S. R. G., 2022. Automated hybrid deep neural network model for fake news identification and classification in social networks. *Integrated SCIENCE TECHNOLOGY*, 10 (2), 110–119. URL `http://pubs.iscience.in/journal/index.php/jist/article/view/1413`.

Santhosh Krishna, B. V., Sharma, S., Devika, K., Sahana, Y., Sharanya, K. N. and Indraja, C., 2022. Review of fake product review detection techniques. *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, 771–776.

Satia Budhi, G., Chiong, R., Wang, Z. and Dhakal, S., 2021. Using a hybrid content-based and behaviour-based featuring approach in a parallel environment to detect fake reviews. *Electronic Commerce Research and Applications*, 47, 101048. URL `https://www.sciencedirect.com/science/article/pii/S156742232100020X`.

Shoryu, T., Wang, L. and Ma, R., 2021. A deep neural network approach using convolutional network and long short term memory for text sentiment classification. *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 763–768.

Sihombing, A. and Fong, A., 2019. Fake review detection on yelp dataset using classification techniques in machine learning. *2019 International Conference on contemporary Computing and Informatics (IC3I)*, 64–68.

Tahvilian, E., Partovi, E., Ejtehadi, M., Bakhshayesh, P. R. and Behzadipour, S., 2022. Accuracy improvement in simple and complex human activity recognition using a cnn-bilstm multi-task deep neural network. *2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, 1–5.

Wang, W. and Gang, J., 2018. Application of convolutional neural network in natural language processing. *2018 International Conference on Information Systems and Computer Aided Education (ICISCAE)*, 64–70.

Weng, Y.-L., Chien, S.-Y. and Lin, S.-Y., 2022. Fake reviews detection with hybrid features using time-sequential deep learning model. *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*, 1–3.

Wu, M. and Tang, Z., 2021. Research on user action recognition method based on parallel cnn-bilstm neural network. *2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, volume 4, 2003–2009.

Zhang, L., Dong, Y. and Wu, Y., 2018. Multi-layer cnn features aggregation for real-time visual tracking. *2018 24th International Conference on Pattern Recognition (ICPR)*, 2404–2409.

Appendix A - Project Proposal

Department of Computing and Informatics

2022-23 Academic Year Individual Masters Project

Project Proposal Form

Please refer to the **Project Handbook Section 4** when completing this form. Note that your proposal should be your own original work and you must cite sources in line with university guidance on **referencing and plagiarism**[1].

Degree Title: **MSc Data Science and Artificial Intelligence**	**Student's Name: SHILPA SHAJI NELLIKKAKUNNEL**
	Supervisor's Name: HARI PANDEY
	Project Title/Area: FAKE PRODUCT REVIEW MONITORING SYSTEM

Section 1: Project Overview

1.1 Problem definition - use <u>one sentence</u> to summarise the problem:

This project focuses on development of a fake product review monitoring system. It will improve reliability of a system especially online purchasing system where user-generated reviews play a central role in selling of a product.

1.2 Project description - briefly explain your project:

The main idea of this project revolves around analysing customer's reviews provided in online shopping sites. Further, this project will be useful to identify and remove the fake reviews and provide actual reviews along with ratings related to the products. To achieve the main goal, a Fake Product Review Monitoring System (FPRMS) is proposed. It will be useful to eliminate the fake reviews and helps in regaining the customer confidence in the system using opinion mining.

1.3 Background - please provide brief background information, e.g., client, problem domain, and make reference to the literature (minimum 4-5 sources):

With development in e-commerce, online shopping takes a significant ratio of the whole market. The reviews are written by people who have bought the product and it has precise details about the products, which can be used as a guide while reviewing potential purchases (Yoshida, Tabata & Hosoda,

[1] https://libguides.bournemouth.ac.uk/study-skills-referencing-plagiarism

Bournemouth University

Department of Computing and Informatics

2022-23 Academic Year Individual Masters Project

2020). The quantity of users, nevertheless, has a significant impact on the review's fineness. Also, the users choose to write brief comments that describe the product's quality and cost (Liu, Fu, Liu & Yue, 2015).
Sellers will occasionally use faked reviews, frequently generated by bots, to raise ratings on their products on e-commerce platforms like Amazon. This can make customers less satisfied with their purchases (Wood & Slhoub, 2022). The importance of identifying spammers and faked reviews increases as the damages caused increases (Li, Feng & Zhang, 2016).

1.4 Aims and objectives – what are the aims and objectives of your project? should be specific and measurable:

The goal is to train four hybrid Machine Learning (ML) models on fake review dataset that could fulfil the objective to predict the accuracy of how genuine the reviews in a are given dataset.

1.5 Research Questions

1) How efficient are the machine learning models in identifying and eliminating the fake reviews and ratings in the marketplace?
2) What will be the most suitable algorithm to fulfil the task of fake review detection and its elimination?

Section 2: Artefact

2.1 What is the artefact that you intend to produce?

A coded script to run machine learning models to predict the accuracy of how genuine the reviews are on the given dataset and produce a comparative study on the four hybrid models.

2.2 How is your artefact actionable (i.e., routes to implementation and exploitation in the technology domain)?

In this project, four distinct types of hybrid models are compared by different techniques like 'Classification + Classification', 'Classification + Clustering', 'Clustering + Classification', and 'Clustering + Clustering', to identify false reviews and eliminate them.

Bournemouth University

Department of Computing and Informatics

2022-23 Academic Year Individual Masters Project

Section 3: Evaluation

3.1 How are you going to evaluate your project artefact?

This artefact will be evaluated and trained using sample data and then after evaluated by using actual data on online shopping sites. There are some real-life experiences which I have come through and the artefacts can be evaluated on those cases to verify the accuracy.

3.2 How does this project relate to your MSc Programme and your degree title outcomes?

The project works on python that involves training of data model using a dataset of reviews.

3.3 What are the risks in this project and how are you going to manage them?

The main risk that comes up is the model selection and the data set selection.
Hybrid model selection for fake review detection brings in a good performance margin rather than the existing models, like Random Forest or Naïve Bayes algorithm.

Section 4: References

4.1 Please provide references if you have used any.

Yoshida, M. Tabata, T. & Hosoda, T., 2020. A Study on Relationship Between Consumer Review Behavior and Purchasing in EC site. International Congress on Advanced Applied Informatics, 9, pp.791-796.

Liu, J. Fu, X. Liu, L, & Yue, K., 2015. What Makes a Good Review: Analyzing Reviews on JD.com. International Conference on Service Science, pp.213-216.

Wood, B. & Slhoub, K., 2022. Detecting Amazon Bot Reviewers Using Unsupervised and Supervised Learning. IEEE World AI IoT Congress, pp.01-08.

Li, Y. Feng, X. & Zhang, S., 2016. Detecting Fake Reviews Utilizing Semantic and Emotion Model. International Conference on Information Science and Control Engineering, 3, pp.317-320.

Section 5: Academic Practice and Ethics

Please delete as appropriate.

Bournemouth University

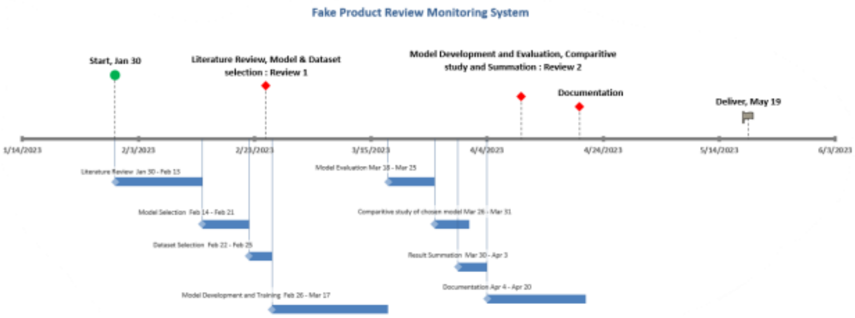
Department of Computing and Informatics

2022-23 Academic Year Individual Masters Project

5.1 Have you made yourself familiar with, and understand, the University guidance on referencing and plagiarism?	**Yes**
5.2 Do you acknowledge that this project proposal is your own work and that it does not contravene any academic offence as specified in the University's regulations?	**Yes**

Note: Please complete the research ethics checklist once the proposal has been approved by your supervisor.

Section 6: Proposed Plan (please attach your Gantt chart below)



Tasks

Start	End	Duration	Label
1/30/2023	2/13/2023	15	Literature Review Jan 30 - Feb 13
2/14/2023	2/21/2023	8	Model Selection Feb 14 - Feb 21
2/22/2023	2/25/2023	4	Dataset Selection Feb 22 - Feb 25
2/26/2023	3/17/2023	20	Model Development and Training Feb 26 - Mar 17
3/18/2023	3/25/2023	8	Model Evaluation Mar 18 - Mar 25
3/26/2023	3/31/2023	6	Comparitive study of chosen model Mar 26 - Mar 31
3/30/2023	4/3/2023	5	Result Summation Mar 30 - Apr 3
4/4/2023	4/20/2023	17	Documentation Apr 4 - Apr 20

Milestones

Date	Label
1/30/2023	Start, Jan 30
2/25/2023	Literature Review, Model & Dataset selection : Review 1
4/10/2023	Model Development and Evaluation, Comparitive study and Summation : Review 2
4/20/2023	Documentation
5/19/2023	Deliver, May 19

4

Appendix B - Ethics Checklist

Bournemouth University

Research Ethics Checklist

About Your Checklist	
Ethics ID	46695
Date Created	05/12/2022 13:49:42
Status	Approved
Date Approved	05/12/2022 16:04:39
Risk	Low

Researcher Details	
Name	Shilpa Shaji Nellikakunnel (s5437538)
Faculty	Faculty of Science & Technology
Status	Postgraduate Taught (Masters, MA, MSc, MBA, LLM)
Course	MSc Data Science & Artificial Intelligence

Project Details	
Title	Fake Product Review Monitoring System
Start Date of Project	30/01/2023
End Date of Project	19/05/2023
Proposed Start Date of Data Collection	22/02/2023
Supervisor	Hari Pandey
Approver	Hari Pandey

Summary - no more than 600 words (including detail on background methodology, sample, outcomes, etc.)

With development in e-commerce, online shopping takes a significant ratio of the whole market. The reviews are written by people who have bought the product and it has precise details about the products, which can be used as a guide while reviewing potential purchases. The quantity of users, nevertheless, has a significant impact on the review's fineness. Also, the users choose to write brief comments that describe the product's quality and cost.

Sellers will occasionally use faked reviews, frequently generated by bots, to raise ratings on their products on e-commerce platforms like Amazon. This can make customers less satisfied with their purchases. The importance of identifying spammers and faked reviews increases as the damages caused increases.

The main idea of this project revolves around analysing customer's reviews provided in online shopping sites. Further, this project will be useful to identify and remove the fake reviews and provide actual reviews along with ratings related to the products.

Filter Question: Is your study solely literature based?

Additional Details

Page 1 of 2 Printed On 16/05/2023 15:23:05

Will you have access to personal data that allows you to identify individuals which is not already in the public domain?	No
Will you have access to confidential corporate or company data (that is not covered by confidentiality terms within an agreement or separate confidentiality agreement)?	No

Storage, Access and Disposal of Research Data	
Once your project completes, will your dataset be added to an appropriate research data repository such as BORDaR, BU's Data Repository?	Yes

Appendix C - First Progress Review Report

Department of Computing and Informatics

Postgraduate Project First Progress Review

To be completed and signed by the Supervisor and Student during week commencing 6 March 2023.

Student: Shilpa Shaji Nellikakunnel	**Supervisor: Hari Pandey**

Assessment

1. Definition of the problem *Has the problem, research aims, and questions been defined, has the artefact been identified and have objectives been set?*	Choose an item. Yes
Comments: The problem is defined. It is mostly complete but still room for further improvement.	
2. Review of literature and related work *Is there evidence of appropriate research?*	Choose an item. Yes
Comments: There are related works for the research. It is mostly complete but critical evaluation and analysis of existing research need to be included.	
3. Methodology and Artefact *Is there evidence of appropriate analysis of the problem and design of a solution and appropriate evaluation?*	Choose an item. Yes
Comments: Working on the models for the study. It is in progress; student is clear about what to be included in this section.	
4. Dissertation *Have sections of the dissertation been written and has the Supervisor seen these?*	Choose an item. Not yet done
Comments:	
5. Planning & Progress *Is there an acceptable plan for this project and is it being followed?*	Choose an item.Yes
Comments: Plan is attached with the proposal and is being followed.	
6. Proposal & Online Ethics Checklist *Are proposal and ethic checklist submitted? Are they approved?*	Choose an item. Yes
7. Overall Assessment	Require minor improvement
Signed: Supervisor: Hari Pandey………………. Student: ……Shilpa……………………............. Date: 15/2/2023	

- Supervisor to retain the signed form and supply the student with a copy if required.
- Supervisor to upload the form on Brightspace and grade as *Satisfactory*, *Requires Major Improvement*, *Requires Minor Improvement*, *Unsatisfactory or Invalid*.

Department of Computing and Informatics

- Supervisor to notify the Project Coordinator if the student is at risk of failing the project or not engaging.

Appendix D - Artefact

Working code is uploaded on GitHub

The artefact contains the following :

- Installation for required libraries
- Download Stopwords from NLTK
- Libraries imported for the functioning of the code
- Load labelled data
- Plot graph for data analysis
- Under Sampling of data
- Split data for classification
- Plot model's training accuracy and validation accuracy
- Plot ROC graph
- Evaluation result normalisation
- Report providing confusion matrix, accuracy and classification report
- Model's accuracy comparison
- Data Exploration
- Data Manipulation
- Data pre-processing
- Definition of random forest
- Definition of multinomial naive bayes
- Definition of convolution neural network
- Definition of long-short term memory
- Definition of CNN-BiLSTM

- Execution of code

The file also contains :

- Dataset used for the study
- Video recording of the execution of code

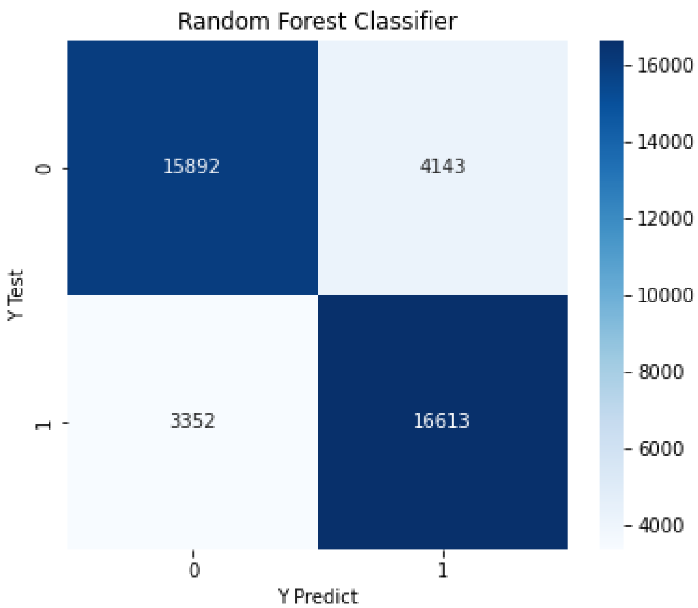
Appendix E - Results

1. Random Forest

Classification Report

```
Random Forest Classifier

Split labeled data into train and test sets
Train a Random Forest Classifier on labeled data
Evaluate the classifier on the test set
Model Evaluation accuracy: 81.2625
Classification Report :
              precision    recall  f1-score   support

           0       0.83      0.79      0.81     20035
           1       0.80      0.83      0.82     19965

    accuracy                           0.81     40000
   macro avg       0.81      0.81      0.81     40000
weighted avg       0.81      0.81      0.81     40000
```

Confusion Matrix

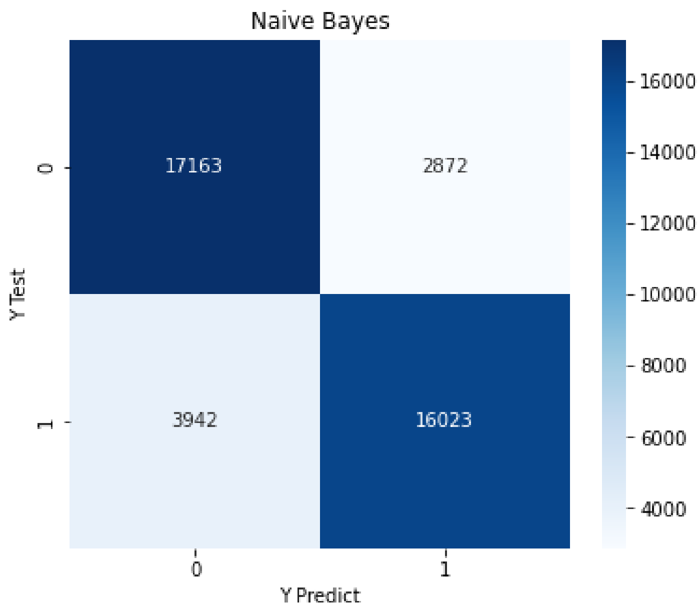


2. Multinomial Naive Bayes

Classification Report

```
Naive Bayes Classifier

Split labeled data into train and test sets
Train Naive Bayes on labeled data
Evaluate the classifier on the test set
Model Evaluation accuracy: 82.965
Classification Report :
              precision    recall  f1-score   support

           0       0.81      0.86      0.83     20035
           1       0.85      0.80      0.82     19965

    accuracy                           0.83     40000
   macro avg       0.83      0.83      0.83     40000
weighted avg       0.83      0.83      0.83     40000
```

Confusion Matrix



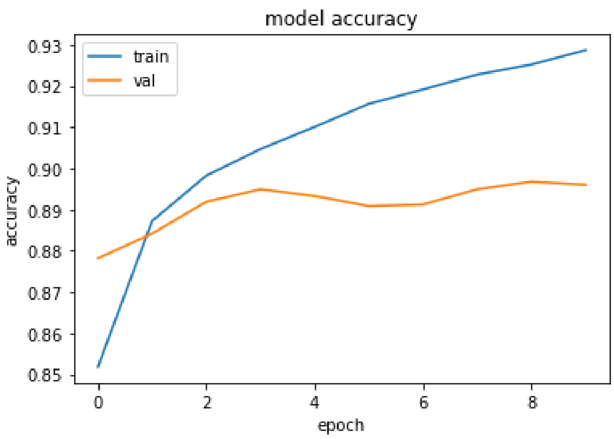
3. Convolution Neural Network

Model Summary

```
Convolution Neural Network

Spliting train data and test data
CNN Model building
Model: "CNN"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 165, 32)           4257440

 dropout (Dropout)           (None, 165, 32)           0

 conv1d (Conv1D)             (None, 161, 64)           10304

 global_max_pooling1d (Globa  (None, 64)               0
 lMaxPooling1D)

 dense (Dense)               (None, 128)               8320

 dropout_1 (Dropout)         (None, 128)               0

 dense_1 (Dense)             (None, 1)                 129

=================================================================
Total params: 4,276,193
Trainable params: 4,276,193
Non-trainable params: 0
```

Epochs

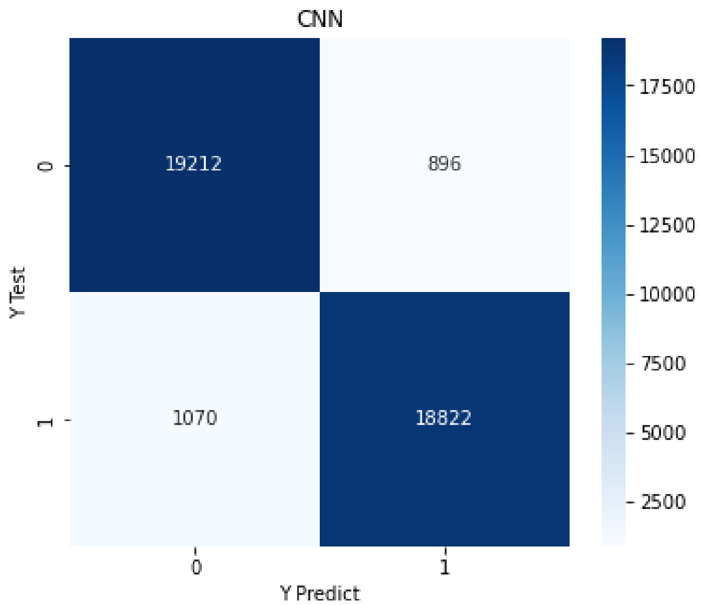
```
Epoch 1/10
4000/4000 [==============================] - 118s 29ms/step - loss: 0.3344 - accuracy: 0.8518 - val_loss: 0.2823 - val_accuracy: 0.8782
Epoch 2/10
4000/4000 [==============================] - 119s 30ms/step - loss: 0.2722 - accuracy: 0.8872 - val_loss: 0.2815 - val_accuracy: 0.8842
Epoch 3/10
4000/4000 [==============================] - 130s 33ms/step - loss: 0.2518 - accuracy: 0.8982 - val_loss: 0.2631 - val_accuracy: 0.8919
Epoch 4/10
4000/4000 [==============================] - 119s 30ms/step - loss: 0.2393 - accuracy: 0.9046 - val_loss: 0.2587 - val_accuracy: 0.8949
Epoch 5/10
4000/4000 [==============================] - 228s 57ms/step - loss: 0.2262 - accuracy: 0.9100 - val_loss: 0.2610 - val_accuracy: 0.8933
Epoch 6/10
4000/4000 [==============================] - 119s 30ms/step - loss: 0.2181 - accuracy: 0.9157 - val_loss: 0.2644 - val_accuracy: 0.8908
Epoch 7/10
4000/4000 [==============================] - 161s 40ms/step - loss: 0.2095 - accuracy: 0.9191 - val_loss: 0.2629 - val_accuracy: 0.8912
Epoch 8/10
4000/4000 [==============================] - 120s 30ms/step - loss: 0.2026 - accuracy: 0.9227 - val_loss: 0.2624 - val_accuracy: 0.8949
Epoch 9/10
4000/4000 [==============================] - 118s 30ms/step - loss: 0.1968 - accuracy: 0.9252 - val_loss: 0.2699 - val_accuracy: 0.8967
Epoch 10/10
4000/4000 [==============================] - 201s 50ms/step - loss: 0.1888 - accuracy: 0.9287 - val_loss: 0.2639 - val_accuracy: 0.8960
Evaluate model on test data
1250/1250 [==============================] - 11s 8ms/step
```



Classification Report

```
Model Evaluation accuracy: 95.085
Classification Report :
              precision    recall  f1-score   support

           0       0.95      0.96      0.95     20108
           1       0.95      0.95      0.95     19892

    accuracy                           0.95     40000
   macro avg       0.95      0.95      0.95     40000
weighted avg       0.95      0.95      0.95     40000
```

Confusion Matrix



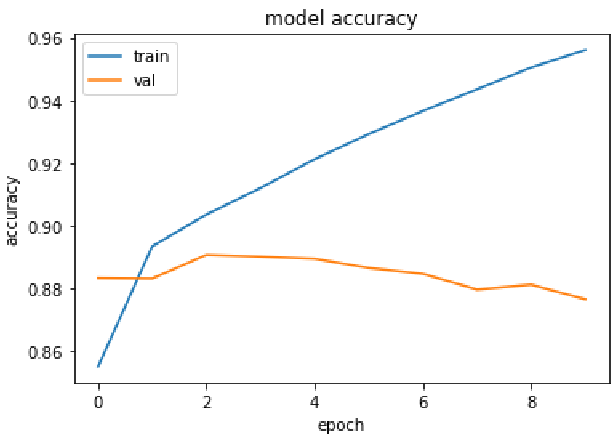
4. Long-Short Term Memory

Model Summary

```
Long-Short Term Memory

Spliting train data and test data
LSTM Model building
Model: "LSTM"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding_1 (Embedding)     (None, 148, 32)           4257440

 lstm (LSTM)                 (None, 100)               53200

 dropout_2 (Dropout)         (None, 100)               0

 dense_2 (Dense)             (None, 20)                2020

 dropout_3 (Dropout)         (None, 20)                0

 dense_3 (Dense)             (None, 1)                 21

=================================================================
Total params: 4,312,681
Trainable params: 4,312,681
Non-trainable params: 0
```

Epochs

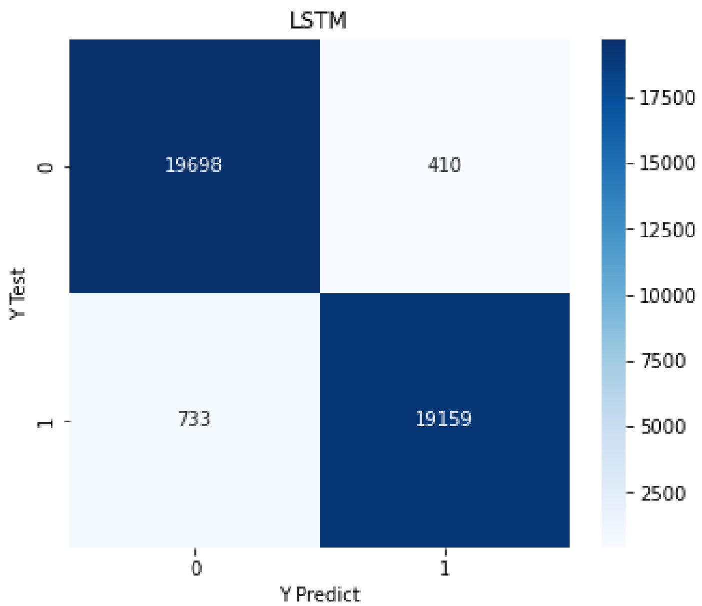
```
Epoch 1/10
8000/8000 [==============================] - 861s 107ms/step - loss: 0.3482 - accuracy: 0.8551 - val_loss: 0.2773 - val_accuracy: 0.8833
Epoch 2/10
8000/8000 [==============================] - 814s 102ms/step - loss: 0.2766 - accuracy: 0.8934 - val_loss: 0.2889 - val_accuracy: 0.8832
Epoch 3/10
8000/8000 [==============================] - 787s 98ms/step - loss: 0.2514 - accuracy: 0.9036 - val_loss: 0.2647 - val_accuracy: 0.8907
Epoch 4/10
8000/8000 [==============================] - 803s 100ms/step - loss: 0.2301 - accuracy: 0.9120 - val_loss: 0.2730 - val_accuracy: 0.8901
Epoch 5/10
8000/8000 [==============================] - 901s 113ms/step - loss: 0.2088 - accuracy: 0.9212 - val_loss: 0.2913 - val_accuracy: 0.8894
Epoch 6/10
8000/8000 [==============================] - 963s 120ms/step - loss: 0.1869 - accuracy: 0.9292 - val_loss: 0.3190 - val_accuracy: 0.8865
Epoch 7/10
8000/8000 [==============================] - 947s 118ms/step - loss: 0.1676 - accuracy: 0.9366 - val_loss: 0.3544 - val_accuracy: 0.8847
Epoch 8/10
8000/8000 [==============================] - 982s 123ms/step - loss: 0.1493 - accuracy: 0.9435 - val_loss: 0.4145 - val_accuracy: 0.8797
Epoch 9/10
8000/8000 [==============================] - 987s 123ms/step - loss: 0.1322 - accuracy: 0.9505 - val_loss: 0.4537 - val_accuracy: 0.8812
Epoch 10/10
8000/8000 [==============================] - 988s 123ms/step - loss: 0.1185 - accuracy: 0.9561 - val_loss: 0.4333 - val_accuracy: 0.8766
Evaluate model on test data
1250/1250 [==============================] - 44s 35ms/step
```



Classification Report

```
Model Evaluation accuracy: 97.1425
Classification Report :
              precision    recall  f1-score   support

           0       0.96      0.98      0.97     20108
           1       0.98      0.96      0.97     19892

    accuracy                           0.97     40000
   macro avg       0.97      0.97      0.97     40000
weighted avg       0.97      0.97      0.97     40000
```

Confusion Matrix



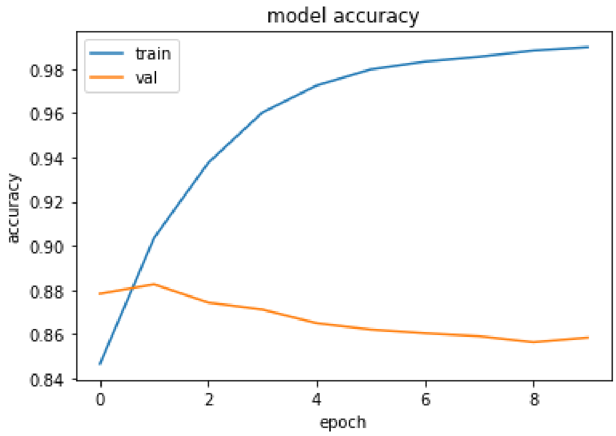
5. **Proposed Model - CNN-BiLSTM**

Model Summary

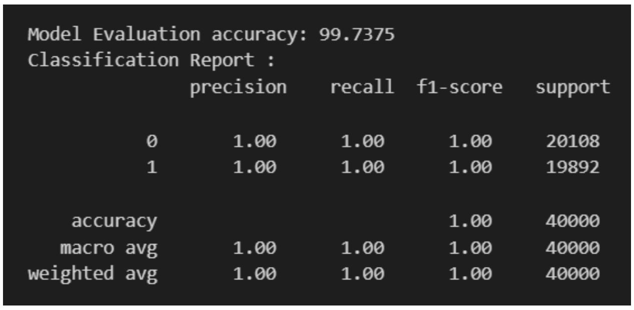
```
Spliting train data and test data
CNN-BiLSTM Model building
Model: "CNN-BiLSTM"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_4 (InputLayer)        [(None, 5000)]            0

 embedding_3 (Embedding)     (None, 5000, 300)         6000000

 conv1d_6 (Conv1D)           (None, 4992, 32)          86432

 max_pooling1d_6 (MaxPooling  (None, 312, 32)          0
 1D)

 dropout_6 (Dropout)         (None, 312, 32)           0

 conv1d_7 (Conv1D)           (None, 306, 32)           7200

 max_pooling1d_7 (MaxPooling  (None, 38, 32)           0
 1D)

 dropout_7 (Dropout)         (None, 38, 32)            0

 bidirectional_3 (Bidirectio  (None, 64)               16640
 nal)

 dense_3 (Dense)             (None, 2)                 130
```

Epochs

```
=================================================================
Total params: 6,110,402
Trainable params: 6,110,402
Non-trainable params: 0
_________________________________________________________________
Epoch 1/10
4000/4000 [==============================] - 9979s 2s/step - loss: 0.3455 - accuracy: 0.8465 - val_loss: 0.2899 - val_accuracy: 0.8783
Epoch 2/10
4000/4000 [==============================] - 9196s 2s/step - loss: 0.2431 - accuracy: 0.9036 - val_loss: 0.2883 - val_accuracy: 0.8826
Epoch 3/10
4000/4000 [==============================] - 9054s 2s/step - loss: 0.1680 - accuracy: 0.9375 - val_loss: 0.3122 - val_accuracy: 0.8743
Epoch 4/10
4000/4000 [==============================] - 7864s 2s/step - loss: 0.1117 - accuracy: 0.9603 - val_loss: 0.3639 - val_accuracy: 0.8711
Epoch 5/10
4000/4000 [==============================] - 6695s 2s/step - loss: 0.0788 - accuracy: 0.9725 - val_loss: 0.4260 - val_accuracy: 0.8650
Epoch 6/10
4000/4000 [==============================] - 6698s 2s/step - loss: 0.0604 - accuracy: 0.9798 - val_loss: 0.4867 - val_accuracy: 0.8621
Epoch 7/10
4000/4000 [==============================] - 7015s 2s/step - loss: 0.0501 - accuracy: 0.9833 - val_loss: 0.5197 - val_accuracy: 0.8604
Epoch 8/10
4000/4000 [==============================] - 6345s 2s/step - loss: 0.0438 - accuracy: 0.9855 - val_loss: 0.5389 - val_accuracy: 0.8590
Epoch 9/10
4000/4000 [==============================] - 6704s 2s/step - loss: 0.0372 - accuracy: 0.9883 - val_loss: 0.5195 - val_accuracy: 0.8563
Epoch 10/10
4000/4000 [==============================] - 7050s 2s/step - loss: 0.0326 - accuracy: 0.9899 - val_loss: 0.5722 - val_accuracy: 0.8584
Evaluate model on test data
1250/1250 [==============================] - 457s 364ms/step
```

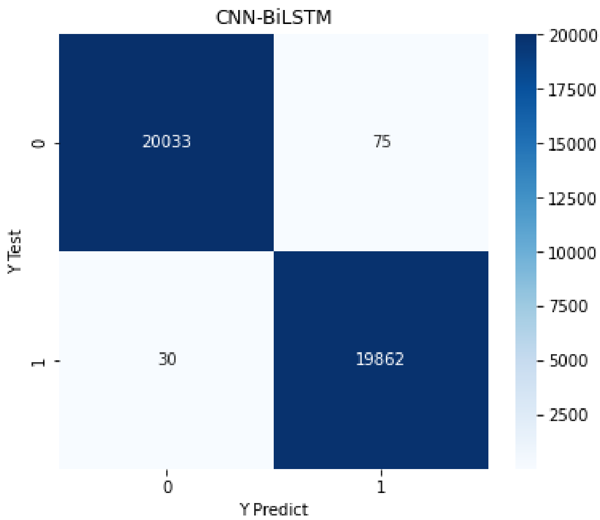


Classification Report

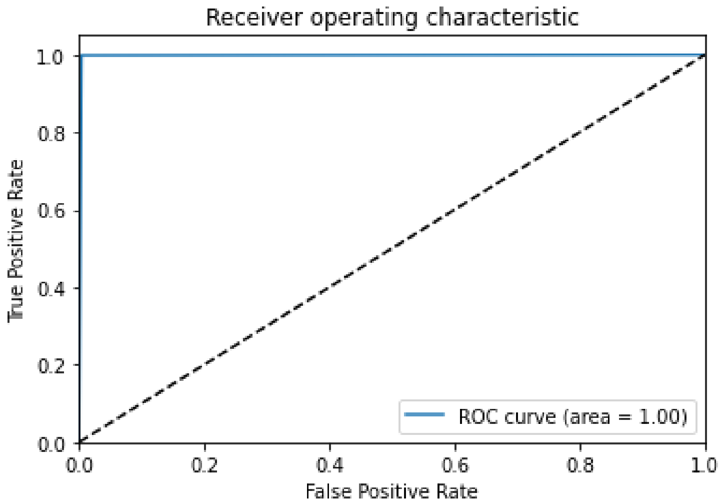


```
Model Evaluation accuracy: 99.7375
Classification Report :
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     20108
           1       1.00      1.00      1.00     19892

    accuracy                           1.00     40000
   macro avg       1.00      1.00      1.00     40000
weighted avg       1.00      1.00      1.00     40000
```

Confusion Matrix



ROC Curve


Receiver operating characteristic
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve (area = 1.00)

Appendix F - Abbreviation

API	Application Programming Interface
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
Bi-LSTM	Bidirectional Long Short Term Memory
C-DSSM	Convolutional Deep Semantic Similarity Model
CNN	Convolutional Neural Networks
CNN-BiLSTM	Convolutional Neural Networks - Bidirectional Long Short Term Memory
DCNN	Deep Convolutional Neural Networks
DL	Deep Learning
FN	False Negative
FND	Fake News Detection
FP	False Positive
GMM	Gaussian Mixture Model
IDF	Inverse Document Frequency
LSTM	Long Short Term Memory
ML	Machine Learning
NB	Naïve Bayes
NLP	Natural Language Processing
RNN	Recurrent Neural Networks
SGD	Stochastic Gradient Descent
SVC	Support Vector Classifier
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TF	Term Frequency
TF-IDF	Term Frequency-Inverse Document Frequency
TN	True Negative
TP	True Positive
XGBoost	eXtreme Gradient Boosting