

Identification of fake reviews using Machine Learning and Deep Learning: A Hybrid Model Approach

Contents

List of Figures	Error! Bookmark not defined.
List of Tables	Error! Bookmark not defined.
Abstract	5
Chapter 1:.....	6
Introduction	6
1.1 Background	6
1.1.1 Ecommerce Businesses	7
1.1.2 Types of Ecommerce	8
1.1.3 Ecommerce Categories	8
1.2 Problem Statement	9
1.3 Research Questions	9
1.4 Research Objectives	9
1.5 SCOPE OF STUDY	10
1.6 Significance of Research.....	10
1.6.1 Significance of customer reviews	10
1.7 Chapter Summary:	12
Chapter 2	13
Literature Review.....	13
Introduction:.....	13
Review of Previous Literature:	13
2.1 Previously Used Datasets and Techniques.....	16
2.2 Effect of Covid on Ecommerce.....	18
2.3 Impact of Fake Reviews.....	19
2.5 Underpinning theory	19
2.5.1 Loss aversion theory	19
2.5.2 Social Exchange Theory:	20
2.5.3 Cognitive Dissonance Theory:.....	21
2.6 Theoretical Framework:.....	21
2.8 Chapter Summary:	22
Chapter 3	23
Research Methodology	23
Introduction:.....	23
3.1 Instruments and Measures.....	24

3.1.1 Fake Review Detection	24
3.1.2 Number of Reviews	25
3.1.3 Number of Positive Reviews.....	25
3.1.4 Reviewer Deviation.....	25
3.1.5 Length of Review	25
3.1.6 Linguistic Inquiry and Word Count (LIWC)	25
3.2 Data Pipeline	25
3.2.1 Data Collection	25
3.2.2 Data Pre-processing	27
3.2.3 Feature Selection:.....	28
3.2.4 Term frequency-inverse document frequency (TF-IDF)	28
3.2.5 Machine Learning Techniques	28
3.3 Classification Models.....	30
3.3.1 SVM.....	30
3.3.2 Extreme Gradient Boosting:.....	31
3.3.3 CatBoost Classifier:	31
3.3.4 Adaboost Classifier:.....	31
3.3.5 Logistic Regression.....	32
3.3.6 Naive Bayes	32
3.3.7 Decision Tree	33
3.3.8 Random Forest	33
3.3.9 K-Nearest Neighbours.....	33
3.3.10 Stochastic Gradient Descent (SGD) Classifier:	34
3.4 Neural Network Models:.....	35
3.4.1 Multilayer Perceptron(MLP):	35
3.4.2 Convolutional Neural Network(CNN):.....	35
3.4.3 Gated Recurrent Unit(GRU):	36
3.5 The Ensemble/Hybrid Model:.....	36
3.6 Chapter Summary:	36
Chapter 4.....	38
Results and Analysis:	38
4.1 Machine Learning Classifiers:	38
4.2 Deep Learning Algorithms:	40
4.3 The Ensemble Model:	42

Chapter 5	44
Conclusion and Future work:	44
Acknowledgement	45
References	46

ANALYZE2GETHER.COM

Abstract

The growing number of fake reviews has become an increasing concern for consumers and businesses in the constantly evolving electronic market. These false reviews may alter consumers' perceptions of goods and services, which may result in poor purchasing decisions and possible harm to businesses. Despite growing research in the field, the challenging issue of reliably recognizing fake reviews still needs to be solved. This study aims to address this gap using a cutting-edge hybrid model strategy that uses deep learning and machine learning capabilities. Identifying and assessing an ensemble model for detecting fake reviews is the primary goal of this study. The ensemble model is a potent tool that enhances overall performance by combining the predictions of various machine-learning models. Multilayer Perceptron (MLP) and Stochastic Gradient Descent (SGD) Classifiers, two distinct classifiers, were used in this work. Both of these classifiers performed well with accuracy ratings of 85.66% and 85.83%, respectively. The ensemble model, which was a result of the combination of these classifiers performed better than its parent algorithms with an accuracy of 86.57%. This increase in accuracy underlines how well ensemble methods work to improve the performance of machine learning models. To ensure the reliability of the ensemble model Cross-validation was used to evaluate its performance. The tiny difference between the cross-validation score and the models accuracy shows that the model is generalizable and performs well on new data. While this study offers a viable solution to the issue of fake review detection, it also paves the way for more research in several areas. By using different ensemble approaches like bagging, boosting, and stacking, there is a window for the improvement of the voting classifier which is the backbone of the ensemble model. Additionally, the performance of the ensemble model can also be increased by tweaking the hyperparameters of the individual. This study's hybrid model, which combines machine learning and deep learning approaches, offers a novel method for identifying fake reviews. The potential of ensemble approaches in addressing complex issues like the identification of fake reviews is highlighted by the proposed ensemble model's higher performance when compared to individual classifiers. This study contributes to the ongoing fight against fake reviews and lays a robust platform for further research in this field.

Chapter 1:

Introduction

In this chapter, our study will focus on figuring out how e-commerce companies could identify fake reviews. We will look at how fake reviews affect how people decide what to buy and how reputable online marketplaces are. The problem statement will show how important it is to have good ways to spot fake reviews to make sure that customer reviews are genuine and reliable. The study will be guided by research questions that aim to understand the characteristics and trends of fake reviews and find effective ways to spot them. Research goals will be laid out to give the study a clear direction. These goals include making and testing machine learning programs or other ways to spot fake reviews. The study's goals will be set, and it will be clear which aspects of identifying fake reviews in the e-commerce business will be examined. Research's importance will be emphasized, focusing on how it could help both consumers and companies make better decisions and keep their trust in online platforms.

1.1 Background

A traditional commerce business involves the practice of buying and selling goods, whereas if this process incorporates any means of the internet, it is termed as E-commerce (electronic commerce). Today many businesses have a physical and online presence to conduct their sales and manage their inventories. According to eMarketer, the online sales at the end of 2022 will break all previous records and will surpass \$5 trillion, and by 2025 it will be \$7 trillion. To understand what e-commerce is, let's take a closer look at its history, growth, and impact on the business world. As there is a continuous advancement on the internet, consumers are becoming more active influencers in online word-of-mouth (WOM) communications. After buying a product they share their experience in the form of a review. According to some American studies, product feedback influences 80% of purchasing decisions (Tufail et al., 2022). The issue is determining if the feedback provided is legitimate or fake. There are two types of reviews, a numeric rating, or a text-based review. Numeric product ratings can reflect how much people appreciate a product, but do not give sufficient information on how a product succeeds or fails. Text reviews, on the other hand, can show users' profound opinions and detailed experiences (Moon et al., 2014). Research shows that 93% of people say online reviews would affect their shopping choices. It shows that most people read online reviews and use them to make decisions (Vimaladevi & Dhanabhakaym, 2012) (Nagi Alsubari et al., 2022) (Barbado et al., 2019). According to a survey, over 60% of customers read online product reviews at least once per week, and 93% think that doing so improves the accuracy of their buying decisions, lowers their risk of losing money, and affects their shopping preferences. 82% of people who do business-to-business and business-to-consumer transactions online read product reviews before buying, and 60% read comments weekly (Chen et al., 2022). Fake reviews are a growing problem in the electronic market, and they could mislead customers and companies while harming their reputations (Baishya et al., 2021).

When a new product is launched, early adopters' product reviews can significantly impact the consumers' judgments, which eventually, decide the commercial success or failure of the product (Moon et al., 2014) (Ahmed et al., 2019). Fake reviews have a significant impact since they directly influence consumers' purchasing decisions. Consumers either reject the product

or purchase it based on their feedback (Tufail et al., 2022) (Yao et al., 2021). The lack of evidence of fakeness in textual reviews makes it difficult to identify fake reviews (Barbado et al., 2019). Because fraudsters use a variety of strategies to make their reviews appear authentic, spotting fake reviews can be challenging (Nagi Alsubari et al., 2022). According to some American studies, online reviews help consumers feel more confident (88% feel this), feel more secure about their purchases (80%), and feel more accountable to their companies (78%). (Pew Research Center, 2020). Customers rely on online reviews, especially when they need more information about a specific product type. Eighty percent of Americans say they look at online ratings and reviews before making a first-time purchase (Pew Research Center 2016) (Watson & Wu, 2021). Additionally, fake reviews can harm or artificially inflate the reputations of businesses (Barbado et al., 2019). In terms of Ecommerce the Word-of-mouth is called eWOM. Positive "eWOM" from happy customers can help boost sales, but brand managers worry that negative "eWOM" will be more noticeable and powerful. When Amazon.com decided in 1995 to let customers post about their experience with the purchased product, some people said it was "a recipe for retail suicide." In contrast, consumer reviews can increase sales, make people less sensitive to price, and make them happier with their purchases by telling them about essential product features and reducing their uncertainty about how these features will work. When people make decisions based on information, they are more likely to be willing to pay a higher price, be satisfied with their choice, and promote the brand online. Practitioners understand that reviews of a product are a part of life and don't try to stop them. Instead, they see them as another part of the marketing strategy that needs to be managed (De Maeyer, 2012). In many countries, posting fake reviews is against the law, and people who do so can get in trouble with the law. Regulatory bodies have acted against businesses and people who post fake reviews. In the United States, for example, the Federal Trade Commission (FTC) has taken action against companies that have been caught putting out fake reviews (Baishya et al., 2021). The detection and prevention of fake news are complex tasks because they require analyzing vast volumes of data from various patterns and features of fake news (Ahmed et al., 2019) (Baishya et al., 2021).

1.1.1 Ecommerce Businesses

The e-commerce businesses are of different sizes, but we can divide them into four major categories:

Table 1 Ecommerce Businesses

Start-up	This type of business is in its initial phases, mostly, an entrepreneur is the one behind the start-up to introduce an innovative business model. A start-up is defined by different criteria. Sometimes it's said that a business is called a start-up if it has less than 100 employees. On the other hand, Alex Wilhelm, the writer of TechCrunch, defines in terms of revenue that if the company has reached the \$50 million revenue threshold, it is no longer called a start-up.
Small Business	A business with less money and fewer employees that is owned by a sole proprietorship, partnership, or corporation falls into the small business category. The US small business administration defines criteria for small businesses to be of 100 to 1500 employees or have an average annual receipt of between \$1 million and \$40 Million.

Mid-Market	Mid-markets are also called SMEs (Small and Medium Sized Enterprises). The business is said to be an SME if the no. of employees ranges from 101 to 500 and the revenue is between \$10 Million and \$1 Billion.
Enterprise	It is the largest type of e-commerce business and has more than 1000 employees and normally generates more than \$1 billion annually.

1.1.2 Types of Ecommerce

Table 2 Types of Ecommerce

B2C	Business-to-consumer is the most common sales model used in e-commerce. In this model, the transactions are made between a business and a consumer. For example, if you buy a product from an online store, it will fall under the Business-to-consumer business model
B2B	Business-to-Business model sales and purchases are made between businesses and do not involve consumers. It is a transaction between a retailer or wholesaler and a manufacturer.
C2C	Consumer-to-consumer is the sale of products and services between customers. It is the earliest form of e-commerce. An example of C2C selling can be seen on eBay.
D2C	Direct-to-consumer is a model in which a business sells its product or services to the consumers directly, without involving any retailer, distributor, or wholesaler. The most common example of D2C can be Netflix, that is serving the consumer directly
C2B	Consumer-to-business is a model which is opposite to the traditional model, in this model, consumers make their products and services available to business buyers. The example of a stock photos website is a great fit for C2B model as the photographs from different photographers is directly available for purchase from the website.
B2A	Business-to-Administration is a type of model that deals with the transactions made between businesses and administration. For example, we can say that products and services deal with legal documents.
C2A	Consumer-to-Administration is like B2A, but in this case, consumers are selling products and services to an administration, for example, online consultations for education or online tax preparation etc.

1.1.3 Ecommerce Categories

Ecommerce can be summarized into selling three major categories.

1.1.3.1 Physical goods

Any tangible product that can be bought and sold in any store or online is said to be a physical good. For example, think of your favorite clothing, home décor, or electronics brand. They all fall under the category of physical goods. These types of eCommerce businesses have a B2C or D2C model.

1.1.3.2 Digital goods

Digital files such as tools, online education, templates, and any other digital product that is downloadable fall under the category of digital goods. Whether you're an online business owner or a small business, digital products can help you sell more online.

1.1.3.3 Services

Providing some specialized services to consumers, such as freelancing in different domains like writing, designing, coaching, marketing, etc., in exchange for compensation, all fall under the category of selling services.

1.2 Problem Statement

Fake reviews are becoming more common on e-commerce platforms, which is a major problem for both customers and companies. These false evaluations have the potential to mislead customers and influence their purchasing decisions, which could result in financial losses and destroy consumer confidence in online platforms. To preserve the trustworthiness and reliability of e-commerce platforms, it is necessary to develop efficient ways for identifying and analysing fake reviews.

Fake reviews are a managerial problem because they can damage a company's reputation and credibility. Operational problems, such as poor customer service or a defective product, can also lead to negative reviews, but these can be more easily addressed and corrected. Fake reviews can be more difficult to identify and remove, and they can have a lasting negative impact on the reputation of a company (Garg et al., 2022).

Existing models do not address this problem in a better way, so the problem that is to be dealt with is to design a model that can address multiple variables accurately (Ahmed et al., 2019), as the models I have studied in the previous literature were trained and tested on a single dataset either Yelp dataset of hotels or restaurants or the dataset obtained from Amazon (Nagi Alsubari et al., 2022).

1.3 Research Questions

RQ1: Are conventional machine learning and deep learning algorithms useful for identifying fake reviews?

RQ2: Can techniques based on ensemble learning increase the accuracy rate for this issue?

1.4 Research Objectives

- Evaluate the effectiveness of different machine learning and deep learning algorithms for detecting fake reviews and examine the limitations and challenges of existing methods
- Identify a hybrid model incorporating machine learning and deep learning algorithms to improve e-commerce platform fake review detection accuracy and efficiency.

Tasks like web mining, feature extraction, neural networks, and machine learning algorithms will be performed for this paper. Addressing this issue is essential because reviews are now a key source of information for customers and aid in their decision-making..

1.5 SCOPE OF STUDY

The main goal of this study is to find and analyze fake reviews on e-commerce sites. To identify and fully understand the traits of fake reviews, it aims to investigate fake reviews, identify algorithms, and develop new techniques. The study's scope will include, among other things, data collection, preprocessing, feature engineering, and the use of machine learning and neural network algorithms. The study will also examine the effectiveness and limits of different approaches, considering review text analysis, sentiment analysis, user behavior patterns, and other relevant factors. The study may also examine how fake reviews affect how people choose products and look for ways to lessen their effects.

1.6 Significance of Research

Consumers will be able to judge a product or service in a better way as there will be no fake reviews that may alter or deviate from consumers decisions, and as a whole, the whole society will be ethical. Also, the platform that has only no fake reviews will be trusted by the consumers, and its business may grow more rapidly.

1.6.1 Significance of customer reviews

Consumers have trouble shopping online because they need more information. They overcome these problems by reading online reviews, which give them information that helps them make better purchases. These reviews help you decide what to buy because they discuss how the product works, not just its features. A study found that 97% of consumers look up information about a product they want to buy online, and 93% of consumers read reviews before making a purchase decision. And most consumers are willing to pay 31% more for the same product if reviews convince them of its positive attributes. So, a product with good online reviews is like one with good electronic word of mouth (eWoM), which affects both current and potential customers' decisions about what to buy. Most e-commerce sites encourage consumers to share their opinions because this is an effective way to get eWoM.

1.6.1.1 Individual

According to a survey, over 60% of customers read online product reviews at least once a week, and 93% of them think that doing so improves the accuracy of their purchasing decisions, lowers their risk of losing money, and influences their shopping preferences (Chen et al., 2022). More than 80% of U.S. consumers check online reviews before making a purchase (Abdulqader et al., 2022). When it comes to online reviews, consumers don't always trust them, so they need to read and compare them carefully. On the one hand, reviews are very helpful because they give consumers important information that helps them choose to spend their money on high-quality products and services (Salminen et al., 2022). Fake reviews can have a significant impact on customer purchase decisions through ecommerce platforms. Fake reviews not only deceive customers into believing they are reading honest opinions, they also give the false impression that certain products are of higher quality than they actually are. This could lead to customers making purchasing decisions based on false information and, ultimately, be disappointed with the product they receive. Fake reviews can be identified by paying close attention to specific features of a reviewer's profile and opinion. Suspiciously high ratings, for example, or ratings that appear too good to be true, should be viewed cautiously. Customers should also be cautious if multiple evaluations appear to have been

submitted by the same individual at various times. Although catching all fake reviews can be tough, shoppers should always be vigilant when making a purchase and do a little research before deciding to purchase. If any bad reviews jump out as probable fakes, a customer should look for more information on the product page or contact the seller directly to clear up any issues or worries. Investing time and effort in research will ensure customers get what they paid for.

1.6.1.1.1 Risk Perception and Uncertainty:

People who read fake reviews feel sceptical and at risk. Customers may find it difficult to distinguish between genuine and fake review with so many ratings, some of which may be fake. Confusion can make it difficult for them to make judgments and make them hesitant to purchase items, particularly if they are expensive or important. By looking at how fake reviews make people feel about risk, businesses can learn how to make people feel better and improve the shopping experience.

1.6.1.1.2 Effects on Consumer Behaviour and Loyalty in the Long Run:

Examining how fake reviews affect customer behavior and loyalty over time is essential. If people keep seeing fake reviews that confuse them or lead them to buy things they don't like, they may lose faith in online shopping. This can cause customers to stop returning and switch to other shopping sites or platforms they think are more trustworthy. Understanding the long-term effects of fake reviews on how people act can help businesses and lawmakers devise good ways to lessen the harmful effects and encourage customers to stick with them.

1.6.1.2 Business

A review can affect the product either positively or negatively. This happens because the algorithms that run online marketplaces use reviews to figure out where a product ranks among others in the same category. So, fake reviews can lead to unfair competition, where the ranking of a product is boosted or reduced artificially (Salminen et al., 2022). This means it is possible to weaponize fake reviews, for example, an unethical firm may manufacture a flood of negative reviews about its competitor. Inundating the market with evaluations of this nature might cause the ranking algorithms used by online platforms to reduce the visibility of the company that is being attacked. When it comes to protecting businesses against unfair competition, it is essential to identify and stop such impacts before they can occur to identify and stop the occurrence of such impacts before they can take place (Salminen et al., 2022).

Fake reviews can seriously harm e-commerce companies by compromising their reputations and causing a lack of market trust. This is particularly true given how strongly online customers rely on reviews when choosing which products to buy. In addition to lowering consumer loyalty, fake reviews can result in lost sales, unsatisfied customers, and reputational damage. Therefore, e-commerce companies must ensure that only honest and genuine reviews are displayed on their site. This can be done by checking the authenticity of reviews, ensuring reviewers aren't receiving discounts or other incentives for submitting reviews and looking at the language used to spot any inconsistencies or indications of manipulation. To spot and delete any false or deceptive content, they must also ensure that their reviews are routinely reviewed and controlled.

1.6.1.2.1 Reliability of E-commerce Platforms:

Fake reviews can make it difficult for individuals to trust and believe in e-commerce platforms. When people read fake reviews, they may start to doubt the truth and reliability of the feedback they use to make smart buying decisions. This loss of trust can hurt e-commerce businesses in many ways, including less customer interaction, lower conversion rates, and, in the end, less money coming in. Understanding how fake reviews affect customer trust is essential if you want to come up with ways to restore confidence and keep shopping fun.

1.6.1.2.2 Changes to the way the market works and unfair competition:

Fake reviews can change how the market works and lead to unfair competition. By giving some goods higher ratings and rankings than they deserve, fake reviews trick customers and give unfair advantages to businesses that do this. This can hurt the ability of honest businesses to compete and upset the market balance, creating an unfair playing field. By looking into how fake reviews change the way the market works, strategies and policies can be made to encourage fair competition and protect both companies and consumers.

1.6.1.2.3 Effect on New and small businesses:

Small and new businesses should pay extra attention to how fake reviews affect them. Positive reviews are a big part of how these businesses build their image and attract new customers. Smaller businesses can be hurt more by fake reviews because they may need more means or a large enough customer base to counteract the adverse effects. Understanding the specific problems that smaller businesses face when fighting fake reviews can help shape strategies and laws to help them grow and stay in business in the e-commerce world.

1.7 Chapter Summary:

This chapter examines e-commerce platform fake reviews. The problem statement emphasizes how fake reviews damage customer trust and company reputation. Existing models cannot accurately detect fake reviews. Hence a multi-variable model is needed. The research questions examine how traditional machine, deep, and ensemble learning methods can improve review detection accuracy. The research goals include analyzing existing methodologies, identifying a hybrid model combining deep learning algorithms and machine learning, and feature extraction and algorithm implementation. The study comprises data analysis, preprocessing, feature engineering, machine learning, and neural network methods. The research helps consumers make informed decisions, builds trust in e-commerce platforms, and provides credible information. Customer reviews influence buying decisions, and tackling fake reviews improves society's ethics and helps businesses grow ethically.

Chapter 2

Literature Review

Introduction:

In this chapter, we will study the literature to learn more about different parts of our research topic. We will look at the used datasets and techniques, the effect of fake reviews, and related theoretical frameworks. The literature study will give us information about the methods used in similar studies, the effects of fake reviews, and the theories that explain how people act when buying things online. By carrying out this thorough review, we hope to lay a solid foundation for our research and better understand the factors that affect e-commerce and the proliferation of fake reviews. This will let us come up with plans and suggestions to improve e-commerce for businesses and customers.

Review of Previous Literature:

The emergence and expansion of e-commerce have greatly aided the development of new digital marketing strategies like product reviews. Customers enjoy making purchases online and rely heavily on product reviews to decide which products to buy. These product reviews are essential bits of information that significantly impact product sales (Garg et al., 2022). Previous studies have mainly concentrated on creating automated techniques, such as sentiment analysis, machine learning, and natural language processing, to identify fake reviews (Nagi Alsubari et al., 2022). Recent advances in artificial intelligence (AI) and business intelligence (BI) based on Big Data Analytics leverage product reviews to understand buyer sentiment. Fake product reviews, however, pose a significant threat to consumer retention. The study aims to analyze recent research on preventing fake product reviews, identify gaps in the literature, and assess if blockchain technology is a viable solution (Garg et al., 2022). An increase in online shopping has been observed worldwide due to the Covid-19 outbreak and the implementation of lockdowns, social isolation, and other preventative measures. The importance of online shopping and the widespread use of e-commerce has created competition among businesses for online sales. Highlights how critical online reviews are in promoting or defaming a company (Ahmed et al., 2019). The importance of product reviews in helping customers make decisions has given rise to a hotly debated subject known as detecting fake or fraudulent reviews. Barbado et al. suggest a feature structure for examining social media sites in the consumer electronics industry. This framework comprises user-centric features that record user behavior on the website and review-centric features that concentrate on the review text (Barbado et al., 2019). Since these ratings affect a company, dishonest people write fake reviews for their own benefit. This study proposed a methodology for detecting fake reviews using text classification and machine learning approaches. The study employed classifiers like Support Vector Machine, K-Nearest Neighbour, and logistic regression (SKL) to identify fake reviews based on the number of pronouns, verbs, and attitudes. On the Yelp and TripAdvisor datasets, the suggested method for finding fake online reviews is more accurate than other cutting-edge methods, with 95% and 89.03% accuracy, respectively (Tufail et al., 2022). But it's important to recognize the limitations of these traditional approaches. When it comes to dealing with high-dimensional data and overcoming the difficulties of data sparsity, deep neural network models clearly outperform standard

methods (Hajek et al., 2020) (Ahmed et al., 2019). Ahmed et al. conducted one study to explore the techniques and challenges of detecting fake news. The study also introduced a new deep learning approach called "FakeDetectorIt" for fake news detection. This proposed model incorporates textual classification and captures the interdependencies among news articles, authors, and subjects. The authors recommend that future research prioritize addressing the challenges related to big data analytics and developing more effective deep learning models for detecting fake news (Ahmed et al., 2019).

Deep learning models have done very well in many areas, such as natural language processing, computer vision, and speech recognition. Deep learning brings potential and challenges in the context of big data analytics, particularly when coping with the large volume, variety, and quality of data (Manzoor et al., 2019). Using artificial neural networks, decision trees, and logistic regression, Fuller et al. investigated different categorization strategies in their study on automated deception detection. On the foundation of information fusion, they also created an ensemble approach. The researchers retrieved 31 text-based deception traits from labeled assertions of crime scenarios. According to their research, the accuracy of the fusion-based model was 74%, while that of the decision tree and artificial neural network models was 73.46% and 71.60%, respectively (Fuller et al., 2011). Online reviews alter how people shop and how businesses react to customer feedback. Online reviews are information flows that affect how marketing systems work at the micro, meso, and macro levels. According to the study, researchers have paid much attention to the effects of online reviews on the flow of information and the complexity of micro- and meso-level efficiency outcomes over the past 20 years. However, the effectiveness-related results of internet reviews need more attention. Online reviews are more than just a resource for businesses or customers. From a macro-marketing point of view, they are a flow of information that can affect the well-being of all stakeholders (Watson & Wu, 2021). Fake reviews portray a bad impression of a product's quality. Therefore, it is essential to find fake reviews. Unfortunately, automatic detection has only partially succeeded in this challenging task. This study showed how to do fake reviews and how to spot them. An Amazon e-commerce dataset tried out two language models, ULMFiT and GPT-2, to do fake product reviews. The better model, GPT-2, was used to make a dataset for a classification task of finding fake reviews. It was shown that a machine classifier could do this task almost perfectly, unlike human reviewers, who do much worse in terms of accuracy and agreement than the tested algorithms. The model was successful in identifying fake reviews that humans created. The results suggest that it can be hard for people to spot fake reviews, but "machines can fight machines" when it comes to doing so. This research helped in consumer protection, protecting businesses from unfair competition, and holding review platforms accountable (Salminen et al., 2022). Deceptive reviews, or fake reviews, have grown in popularity. Online marketing transactions like selling and buying are rising rapidly. Customers can leave reviews on e-commerce sites. Before buying, new buyers read website reviews and comments. However, new users need to learn to distinguish between genuine and fake reviews, which trick customers, cost firms money and hurt their reputations. This research develops an intelligent approach to detecting fake reviews on eCommerce platforms using n-grams of review text and reviewer sentiment scores. This work employed a common fake hotel review dataset for experimentation and data preparation using a term frequency-Inverse document frequency (TF-IDF) strategy for feature extraction and representation. The models used n-grams of review texts to find fake or real reviews and

classify them. The Trip Advisor dataset was used to train and test four supervised machine-learning algorithms. Based on testing accuracy and F1-score, naïve Bayes (NB), SVM, AB, and RF received 88%, 93%, 94%, and 95%, respectively. The proposed methods outperformed the accuracy of comparable methods on the same dataset (Nagi Alsubari et al., 2022). Online reviews are more important for reputation management as e-commerce platforms evolve. They also influence end-user decisions. A good review for a product usually boosts sales. Today, fake reviews are made to develop an online reputation and attract buyers. As a result, detecting false reviews is an active research topic. Identifying fake reviews depends on reviewer behavior and other crucial features. This research proposes machine learning to detect fake reviews. This research also uses feature engineering to extract reviewer behaviors. The research analyzes many experiments on an actual Yelp dataset of restaurant reviews with and without user behavior variables. KNN, Naive Bayes (NB), SVM, Logistic Regression, and Random Forest are compared in both scenarios. Evaluations also consider the bi-gram and tri-gram language models of the n-gram. KNN(K=7) has the best f-score (82.40%). The extracted reviewers' behavioral traits enhanced the f-score by 3.80% (Elmogy et al., 2021). E-commerce is the most rapidly expanding sector of the economy. Consumers rely heavily on online reviews to evaluate and compare products and services. As a result, more fake reviews (called "opinion spam") hurt customers and service providers. It's hard to find opinion spammers automatically for several reasons, such as the need for well-labeled data. This constraint limits the use of an off-the-shelf machine learning algorithm. This research introduces a new way to tell if reviewers are spam by combining machine learning with a message-passing method that uses the graph structure of the users to make up for the possible lack of labeled data. Instead of always picking the same label for the training step, a new "active learning" method is used. Experiments on three massive real-world datasets from Yelp.com show that this strategy beats active learning and machine learning methods that need a much bigger labeled data set for training (Danilchenko et al., 2022).

An eye-tracking study examined how internet product reviews affect consumer purchases. The study created a conceptual framework for online product reviews and buying intentions based on gender, visual attention in comments, and (ii) empirical research into consumers' fixations during the purchase decision process and behavioral analysis. Consumers paid more attention to negative reviews than favorable ones, especially women. The study also found a strong link between visual browsing and purchase intent. Consumers could not spot fake reviews either. The new study shows in detail how online reviews affect shopping behavior. It shows for the first time how gender affects shopping behavior and explains it from the point of view of attentional bias, which is essential for the theory of how people shop online. Female shoppers pay much more attention to negative reviews than favorable ones. These findings suggest practitioners should prioritize negative reviews and resolve them quickly by customizing product/service information to consumer attributes, including gender (Chen et al., 2022). Online customer reviews, sometimes called "electronic word of mouth," have become very useful in e-commerce because they help people decide what to buy next time. This study examines the primary and secondary factors affecting a review's usefulness. These factors include what the review is about, how many stars it has, how long it is, and how many votes it has. It looks at reviews on Amazon.com for video games, digital music, and grocery store items. The study finds the polarity of feelings, finds patterns of feelings, and then looks at how valuable reviews are under the moderating effect using the text mining approach. The analysis

concludes that the central route has no substantial impact on search goods. According to the study, peripheral sources substantially impact search results. Our research sheds light on how online businesses might develop product-specific marketing strategies (Guha Majumder et al., 2022). Consumers' purchase decisions are affected by online reviews. However, it is still hard to automatically spot fake reviews on the internet, and the methods we have now need to be better to stop the growth of fake reviews. This study defined the chosen constructs using verbal and non-verbal attributes to test the suggested model empirically. Then, we took features from the Yelp datasets and trained four machine learning algorithms—Logistic Regression, Naive Bayes, Decision Tree, and Random Forest—using those attributes. The study showed that the most critical ways to spot fake reviews are number, not being recent, emotion, informality, consistency, the credibility of the source, and differences in behaviour. The study found that non-verbal factors are more important than verbal ones and that combining features from both types improves prediction performance. This theory-based approach made the results easy to understand and not too complicated. It was better than most current methods for finding fake reviews (Abdulqader et al., 2022). Another method to find fake reviews that uses a neural network-based model and aspect-level analysis has been suggested. This method considers conversation and document-level semantics while extracting aspects from reviews. This makes the process of computing less complicated. The writers did tests to see how well their method worked and compared it to other already-used methods (Bathla et al., 2022).

2.1 Previously Used Datasets and Techniques

The detection of fake reviews extensively uses machine learning techniques, with supervised learning as the primary methodology. Supervised learning and unsupervised learning are two subcategories of machine learning. Supervised learning is the most widely used technique for spotting fake reviews. Supervised learning develops models that reliably categorize reviews as real or fake using labeled data. This method uses patterns and traits discovered from labeled data to build assumptions about unread reviews. Algorithms for supervised learning are essential in fake review identification since they can quickly spot fraudulent strategies (Martinez-Torres & Toral, 2019) (Cardoso et al., 2018). Yelp's review dataset consisting of 103,020 recommended reviews and 32,393 not-recommended reviews, was collected and processed in an experiment. The impacts of reviewers' social interactions on fraud detection in online consumer reviews were studied. Afterward, the social interaction and behavioural features were extracted, and the backpropagation neural network algorithm was implemented to predict which review was genuine and which was fake (Goswami et al., 2017). Opinion spam in reviews was studied, and three types of spam were identified. The first was to identify duplicate reviews. Type 2 and type 3 spam reviews were then identified using supervised learning, in which training data was labelled manually. Type 1 was difficult to detect as the data for type 1 spam had to be labelled manually, so duplicate spam reviews were used as positive training data and other reviews as negative training data to create a model. The logistic regression model proved to be very effective in this case. The experiment was done on Amazon reviews, and the results were 78% in terms of area under the curves (Jindal & Liu, 2008). In another research, a real-life Yelp product reviews data set was used to detect fake reviews by using the SVM model. The dataset contained 5678 reviews and 5,124 reviewers from hotels and 58,517 reviews and 35,593 reviewers from restaurants. In this study, two types of features were extracted. The linguistic features were composed of n-grams and the parts of

speech(POS), and the reviewer's features consisted of LIWC and behavioural features. Kullback–Leibler divergence (KL) method was used to differentiate between true and fake reviews, and the accuracy obtained was 84% and 86% for linguistic features and reviewer features, respectively. Features focusing on the user are more effective at spotting fake reviews. While other aspects also help to increase overall accuracy, these offer information about unusual reviewing activities (Barbado et al., 2019).

In the area of detecting fake reviews, deep neural network models including Deep Feedforward Neural Networks (DFFNNs), Convolutional Neural Networks (CNNs), and General Regression Neural Networks have become increasingly popular. These models can efficiently handle the issues of data sparsity and high dimensionality by extracting nuanced characteristics from high-dimensional data (Hajek et al., 2020). The Fake Feature Framework (F3), a feature framework created to identify fake consumer electronics reviews, is introduced in A Deep Learning Approach. F3 combines user-centric and review-centric features, gathering data from the review content and the user's online activity. The authors show that user-centric features—especially those connected to reviewing activity—are superior to review-centric characteristics alone in detecting fake reviews. The authors suggest a deep learning method that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to categorize reviews as fake or real. This method outperforms the conventional machine learning algorithm's ability to detect fake reviews with high accuracy. The study highlights the need to take into account user-centric variables in the identification process. It offers insightful information about the implementation of deep learning approaches for fake review detection (Barbado et al., 2019). By adding numerous behavioral factors, such as the timing and frequency of review submissions, the SAFER model seeks to improve performance (Baishya et al., 2021). The SAFER model uses deep learning and neural networks to improve accuracy in identifying fake reviews, in contrast to conventional systems that mainly rely on manually created features. The extraction of complex characteristics from high-dimensional word, sentence, and emotion representations is made possible by the use of Deep Feedforward Neural Networks (DFFNNs) and Convolutional Neural Networks (CNNs). The usefulness of these suggested models is demonstrated by experimental findings on four real-world fake review datasets. Surprisingly, the suggested models outperform current baseline approaches and cutting-edge fake review detection techniques, as shown by gains in accuracy, AUC, and F-score. In this study, a continuous bag of words (CBOW) and skip-gram methods were used which are parts of NLP. A CBOW architecture tries to understand the context of the sentence and takes that as input to predict the words that are accurate contextually. Skip-gram is an unsupervised learning technique and is used to find the most related words for the input word. This paper used four different algorithms, i.e., Naïve Bayes, Support Vector Machines, LR, and Random Forest for the classification of customer reviews. The experiment showed RF in combination with CBOW, achieved a maximum efficiency of 91% (Bansal & Srivastava, 2018). A linear SVM model is used for the detection of the fake reviews in this article which was based on n-gram features. A standard fake hotel dataset was used for evaluation purposes and was collected by tripadvisor.com. The TF-IDF (term frequency-inverse document frequency) method resulted in a 90% accuracy of the model (Ahmed et al., 2017).

This study also used multiple machine-learning models that are Support Vector Machines, Naïve Bayes, and Maximum Entropy (ME), for the classification of positive, negative, and neutral reviews. 4 and 5 stars are said to be positive reviews, 1 and 2 stars to be negative reviews, and 3 stars as neutral, which is the most difficult to identify whether it is positively neutral or negatively. Machine learning models were trained on unigrams and weighted unigram models and the SVM model gave the highest accuracy, which was 81%. In this article, the binomial regression technique was used to identify fake reviews having abnormal attribution ratings that deviated from public opinion. The authors highlight the value of sentiment analysis while reviewing the body of previous literature on fake review detection. They talk about the sentiment analysis technique, commonly called opinion mining, and how it applies to actual movie reviews. We compare the performance of four supervised machine learning algorithms: Naive Bayes, Support Vector Machine, K-Nearest Neighbors, and Decision Tree (Baishya et al., 2021). Opinion spammers were analysed and detected based on the rating deviation of the products. In this method, honest reviewers' opinions were compared with the fake ratings based on which each reviewer's honesty was calculated (Savage et al., 2015). Various behavioral features, including review timing and submission frequency, are incorporated into the SAFER model based on deep learning models. Its goal is to improve fake review detection accuracy compared to conventional models and clustering methods. The study begins with gathering an Amazon Product and Review Dataset that spans 1996 through 2014. Following the division of this dataset into training and testing sets, preprocessing techniques are used to remove superfluous characters and words (Baishya et al., 2021).

2.2 Effect of Covid on Ecommerce

The world is constantly on a trajectory of digitization. In addition, e-commerce is on the rise. After Covid, online shopping has gained a boost. People are readily turning to the online market instead of entering physical spaces. The convenience is brought right to them.

The recent pandemic covid19 also affected eCommerce and changed the nature of the business as more and more consumers were bending towards online shopping. According to research, 52% of consumers were avoiding traditional commerce and crowded areas. The effects of the pandemic were different on different natures of products. For example, food, groceries, and healthcare had a drastic boost in their sales, but on the contrary to this, products like shoes didn't have a great impact on them, so many manufacturers shifted their business models as some shoe manufacturers started making face masks or some spirit manufacturer started making hand sanitizers, etc (Bhatti et al., 2020).

Table 3 Effect of Covid19 on Ecommerce

Sr. No.	Websites	Revenue millions \$
1	Amazon.com	4059
2	Ebay.com	1227
3	Rakuten.co.jp	804
4	Samsung.com	648
5	Walmart.com	614
6	Appel.com	562
7	Aliexpress.com	532
8	Etsy.com	395

9	Homedepot.com	292
10	Allegro.pl	272

(Andrienko, 2020)

2.3 Impact of Fake Reviews

Reviews can provide valuable insights into customer expectations, preferences, and needs. They can also help businesses to identify areas where they may need to make improvements. Additionally, online reviews can help in building trust and credibility with potential customers. So, fake reviews are bad for a business because they can damage its reputation and credibility. Customers may be less likely to use the business if they see that it has fake reviews, and other businesses may be less likely to want to partner with or recommend the business (He et al., 2021). Fake reviews can be used to manipulate public opinion about a product or service. This can lead to people making poor decisions about what to buy or using a service that isn't right for them. Second, fake reviews can impact a company's bottom line. If people believe that a product is better than it is, they may be less likely to buy it from a competing company. Finally, fake reviews can damage the credibility of online review sites. If people can't trust that the reviews, they're reading are real, they may be less likely to use the site in the future (Salminen et al., 2022).

As the space is getting crowded because more businesses are now working on their e-commerce fronts, it is difficult to distinguish the brands that provide superior quality from those that make fake promises. Moreover, the purchase decision can become even more complicated if the reviews that are left by the customers are fake. Since people cannot physically see or feel the product, they rely heavily on previous customers' words. Fake reviews are becoming extremely common. Businesses can engage in unethical practices and leave positive reviews just to increase sales. In other cases, competitors can post negative reviews to discourage potential clients from making a purchase (Nagi Alsubari et al., 2022). The above cases can affect the customer experience greatly. Positive reviews lure customers into buying more of that product or from that brand, but on the other hand, negative reviews can make decrease sales in e-commerce. Therefore, merchants are dependent on the feedback provided by customers to continuously increase the quality of products. So, it is essential to have a model that can detect fake reviews. The paper will use machine learning techniques.

2.5 Underpinning theory

Below of the underpinning theory for this thesis:

2.5.1 Loss aversion theory

The Loss Aversion Theory, suggested by Daniel Kahneman and Amos Tversky, tells us much about how people feel about potential losses versus potential gains. This theory is related to the idea of fake reviews, and it can help explain why people might do such dishonest things. Loss Aversion Theory says that people tend to have a stronger emotional response when they think they might lose something than when they think they might win the same amount. Such a tendency to avoid losses can be observed in cases of fake reviews, where individuals may

post deceptive opinions to shield themselves or others from potential losses or adverse outcomes.

When business owners receive negative reviews, the fear of losing consumers or gaining a bad reputation can motivate them to post fake positive reviews. They perceive the negative reviews as potential financial and business losses. By publishing fake positive reviews, they try to offset the negative comments and maintain a positive reputation to avoid losing money. People with negative experiences with goods or services may be driven to post fake negative reviews. They may feel obligated to inform others so they do not experience the same loss or disappointment. They aim to prevent others from suffering financial loss or negative feelings by making their negative experience appear worse than it actually was. Researchers and practitioners can better understand the biases and cause behind people writing fake reviews by combining the problem of identifying fake reviews with the Loss Aversion Theory. This data can be used to develop more effective identification techniques and algorithms that account for the affective responses associated with loss aversion. Detection algorithms can be designed to identify patterns in review content that indicate a possible bias toward averting losses or exaggerating negative experiences. For instance, algorithms can analyze the language used to write a review to identify emotional cues or differences that may indicate a fake review. Also, understanding how loss aversion operates can help experts devise methods for combating the biases and emotional responses that can result in fake reviews. To solve the issue of fake reviews, it is crucial to educate customers on the potential biases and motivations behind such dishonesty. Consumers will be less impacted by fake criticism if they are aware of the existence of fake reviews and are provided with tools to determine the veracity of online reviews. The Loss Aversion Theory, in conclusion, offers significant insight into how people perceive possible losses, which can help us comprehend why people leave fake reviews. Researchers and practitioners may develop more effective methods to identify and combat fake reviews by applying this theory to the issue of detecting fake reviews, ultimately resulting in a more reliable online review ecosystem.

2.5.2 Social Exchange Theory:

Social Exchange Theory is a sociological concept that explains why people engage in dishonest behavior, such as writing fake reviews. When individuals interact with one another, they generally do so with the expectation of receiving something in return. People may create fake reviews for various reasons, including financial gain, personal gain, or the urge to obtain a competitive advantage. One of the primary motivations for writing fake reviews is the possibility of financial gain. People who write positive reviews of a business's products or services may receive monetary compensation or a discount. By doing so, these individuals hope to influence consumer decisions, increase sales, and improve the overall appearance of the product or service. Fake reviews may additionally be an option to generate income for the author. People can write good reviews for their businesses or businesses owned by friends or family. By making people feel good about the business, they hope to get more users and make more money. People can also be pushed to write fake reviews by the desire to win a competition. When there is a lot of competition in a market, businesses may do things that aren't ethical, like post fake negative reviews about their rivals' products. This strategy is used to hurt the reputation of businesses competing with them and get customers to buy their goods or services instead. Social Exchange Theory can be used to find fake reviews by figuring out

what drives people to lie in this way. Researchers and practitioners can look at the rewards and incentives that come with fake reviews to find trends and come up with ways to stop and find fraud. Businesses can take steps to stop people from writing fake reviews by finding out why they do it. This could mean enforcing stricter verification processes for online reviews, monitoring suspicious activity on online platforms, and teaching customers about the risks of depending only on online reviews.

2.5.3 Cognitive Dissonance Theory:

Cognitive dissonance theory explains what happens to a person's mind when their views or attitudes are at odds with what they do. Regarding fake reviews, people who post false opinions on purpose may feel cognitive conflict because their actions go against what they believe or value about honesty and integrity. People often explain their actions as a way to reach personal or organizational goals to get rid of the discomfort caused by cognitive dissonance. For example, someone might tell themselves that they need to write a fake positive review to help their business or a friend succeed. Understanding cognitive dissonance is important for dealing with the problem of fake reviews because it shows how people may feel conflicted inside when they do things that aren't true. Researchers and practitioners can come up with ways to reduce cognitive confusion and stop people from writing fake reviews if they understand the psychological processes at play. One way to deal with cognitive dissonance is to encourage people to act ethically and build a culture of honesty and openness. This can be done with the help of training programs, ethical guidelines, and creating a place where people feel safe talking about unethical practices. Also, providing individuals with alternatives to the use of fake reviews may reduce cognitive dissonance. For instance, businesses may offer real incentives, such as customer loyalty programs or referral bonuses, to obtain genuine consumer feedback and suggestions. Social Exchange Theory and Cognitive Dissonance Theory provide crucial insight into the causes and mental processes resulting from fake reviews. By understanding these theories, researchers and practitioners can devise effective strategies for detecting and preventing fake reviews, promoting ethical behavior, and making the Internet safer for online shopping.

2.6 Theoretical Framework:

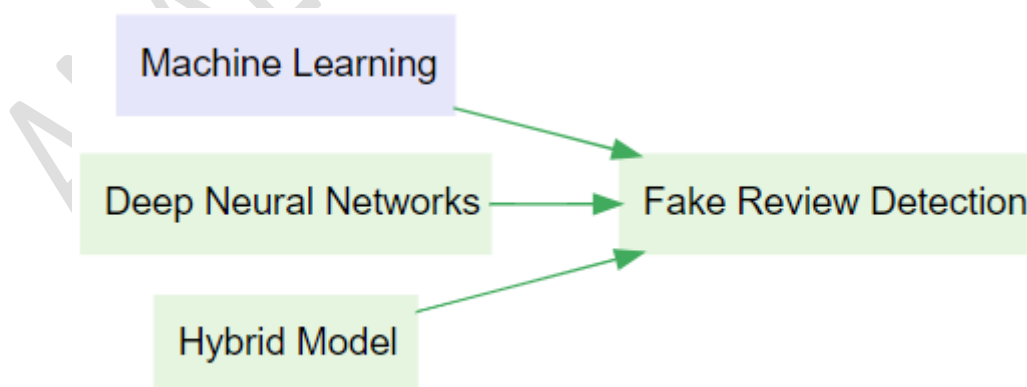


Figure 1 Framework

In this research our aim is to study the methods for the detection of fake reviews. We will be considering the detection variable as or DV and and to study which method is best for this

purpose will be IV's for example in this case we are focusing on whether machine learning will be best for the detection of fake reviews or deep neural networks are better approach or the hybrid model that is proposed to be of a combination of a Machine learning and deep learning gives better results than the parent node.

H1: Machine learning is the best approach for the detection of fake reviews.

H2: Deep neural networks are best for detection of fake reviews.

H3: A Hybrid/Ensemble of these is best fit for the problem of detection of fake reviews.

2.8 Chapter Summary:

In this chapter we studied comprehensively the previous literature and the data set and techniques used previously. We took a comprehensive look at the research that has been done on the topic of e-commerce. We analyzed the influence that Covid-19 had on the industry, investigated the impact of fake reviews on consumer behavior, and built a theoretical framework to direct our research. We also looked at datasets and approaches that have been used in the past. We will be able to address our research questions and contribute to the existing body of knowledge in the field of e-commerce due to this exhaustive analysis of the current information and ideas, which will serve as a firm foundation for our study and form the basis of our investigation.

Chapter 3

Research Methodology

Introduction:

In this chapter, we will look more closely at the techniques used for the study. First, we will provide an overview of the instruments and measures utilized throughout the data collection and analysis processes. We will outline the data pipeline, which includes collecting data, preprocessing the collected data, and selecting features. Techniques for data preprocessing will be implemented to clean and prepare the data, and methods for feature selection will be applied to determine the features that are most relevant for the identification of fake reviews. In order to accurately capture the textual data, we will use a technique known as the term frequency-inverse document frequency (TF-IDF). In addition, we will investigate the machine learning and deep learning strategies that were applied in the research. By providing a detailed account of the research methodology, the aim is to ensure transparency and reproducibility and contribute to the overall validity and reliability of the study.

There are numerous methods for detecting fake reviews on e-commerce platforms. One approach is using machine learning to construct an algorithm to identify fake reviews. This approach can be efficient because it can consider multiple indicators of a fake review, such as the review's text, the user's history, and the overall sentiment of the product's reviews (Baishya et al., 2021).

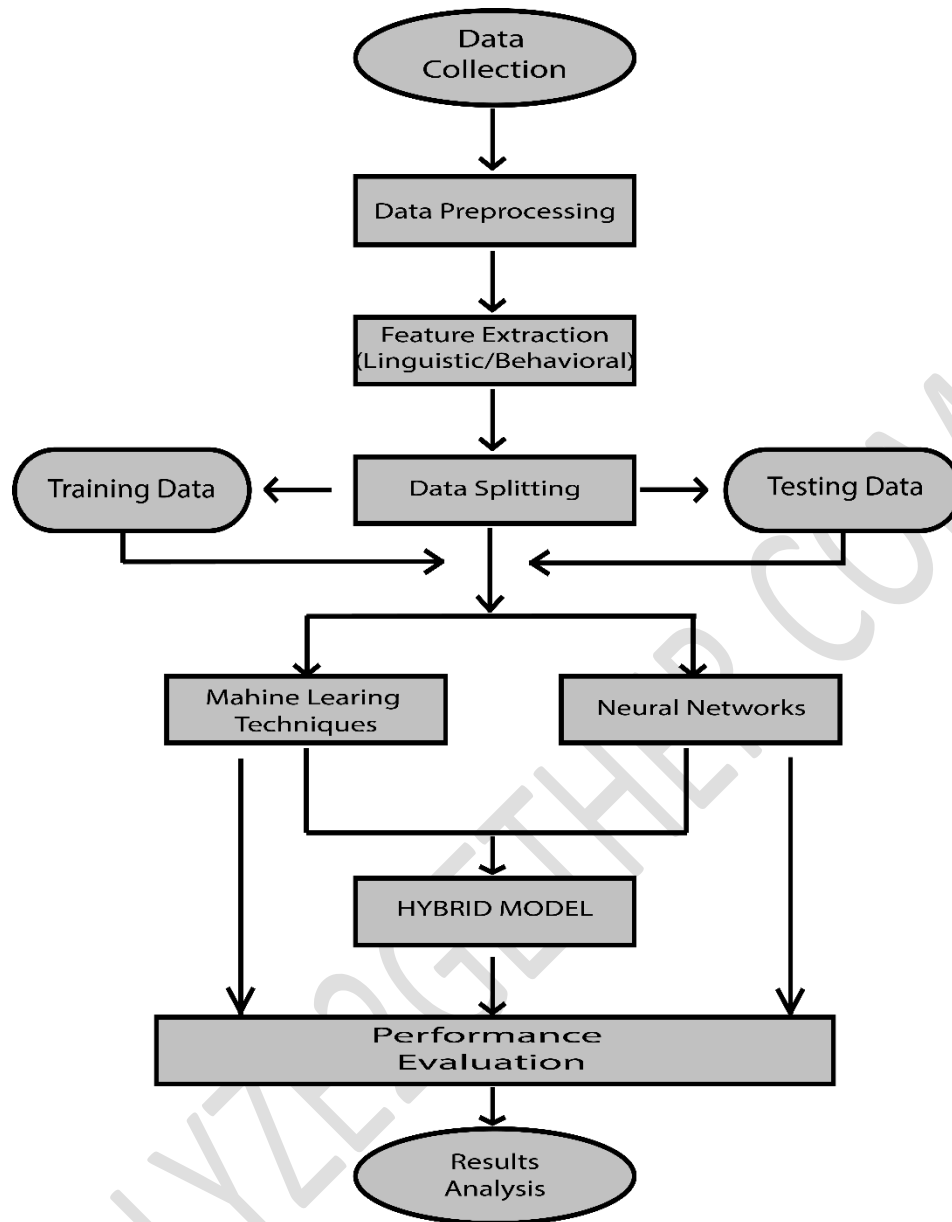


Figure 2 Data Pipeline

3.1 Instruments and Measures

3.1.1 Fake Review Detection

It is a natural language processing method that analyzes, detects, and categorizes product reviews as genuine or fake. Many studies have been published on fake review detection for

some years as it has a significant effect on eCommerce businesses. It is classified into two main features, i.e., Behavioral features and Linguistic features (Nagi Alsubari et al., 2022).

3.1.2 Number of Reviews

In some studies, it is observed that a true reviewer can write one review per day on a service or product purchased. More than 70% of review spammers were noticed to be giving more than five reviews per day, so based on no. of reviews, some fraudsters can also be filtered. (Heydari, Tavakoli, Salim, & Heydari, 2015), (Nagi Alsubari et al., 2022)

3.1.3 Number of Positive Reviews

It is seen that more than 80% of the spam reviewers had a high percentage of positive reviews (according to the article, more than 85% of the reviews are positive). This technique can also filter out most of the fake reviewers (Crawford et al., 2015).

3.1.4 Reviewer Deviation

The normal honest rating of a truthful reviewer is different from a fake reviewer. This technique can also be used to identify review spammers by comparing honest reviewers' opinions with fake ones.

3.1.5 Length of Review

It is observed that the length of the review text of the fraudulent review is short as compared to the original one, according to a research 75% of fake reviewers can only write a maximum of 136 words per review, and when compared to honest reviewers, 90% of the reviewers' words per review were noticed to be an average of 200 words (Noekhah et al., 2014).

3.1.6 Linguistic Inquiry and Word Count (LIWC)

This program analyzes text files and calculates each category's total number of words. James W. Pennebaker and his associates at the University of Texas at Austin created it. LIWC can assist in detecting fake reviews by analyzing the review's overall linguistic style and specific words. For instance, if a review is highly positive but lacks specifics, it may be more probable to be fake. In addition, LIWC can identify specific words that frequently appear in fake reviews, such as "amazing" and "fantastic." By analyzing the overall style and individual words, LIWC can help detect fake reviews.

3.2 Data Pipeline

3.2.1 Data Collection

Our study looked at how well different machine learning algorithms and neural network architectures could find fake reviews on e-commerce websites. We used a dataset with almost 20,000 fake reviews and 20,000 real reviews which can be seen below in the figure 2 from a previous study (Salminen et al., 2022). Salminen et al. made the dataset public to help in the advancement of better models for identifying fake reviews. The characteristics of the data are explained in the table 4 given below in which we can see that the total number of the reviews is 40432 in which 50% of them are labeled as fake and the other half as original reviews that are written by humans. The machine learning methods we used came from the scikit-learn

(SKL) library. They were Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Linear Regression (LR), AdaBoost, and CatBoost. We also used different kinds of neural network architectures, such as Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Gated Recurrent Unit (GRU). By comparing how well these Machine Learning algorithms, neural networks, and a combination of these models worked, we hoped to find the best way to spot fake reviews on e-commerce sites.

Table 4 Data Characteristics

Metric	Value
Total number of reviews	40433
Number of fake reviews	20216
Number of original reviews	20216
Label	1
Longest review length	1462
Shortest review length	0
Maximum number of words in a review	273
Minimum number of words in a review	0
Average number words in a review	35.83
Number of reviews with zero words	2

The review's contents are in text format. Based on this feature, the primary analytical tasks such as sentiment analysis and linguistic features were employed to obtain a textual feature. The sentiment score is a method for figuring out whether a text is positive, negative, or neutral and for computing its polarity score. When reviews reflect a substantial amount of negative sentiment, negative scammers are typically used to using more negative phrases than good ones. The emotion score should be calculated for each review text because positive fraudsters always have a habit of using more positive terms. The next phase was using POS features to check the length of the reviews. POS tagging is the process of assigning a POS tag to each word in the text's content depending on where it appears and how it is used in the sentence. The number of adjectives, nouns, prepositions, coordinating conjunctions, determiners, pronouns, verbs, predetermines, and adverbs were extracted using this technique from the review text. This section concluded that real reviews had more nouns and adjectives whereas fraudulent reviews contained more verbs and adverbs. N-gram features are the results of choosing 'N' adjacent words as the features from the text's content.

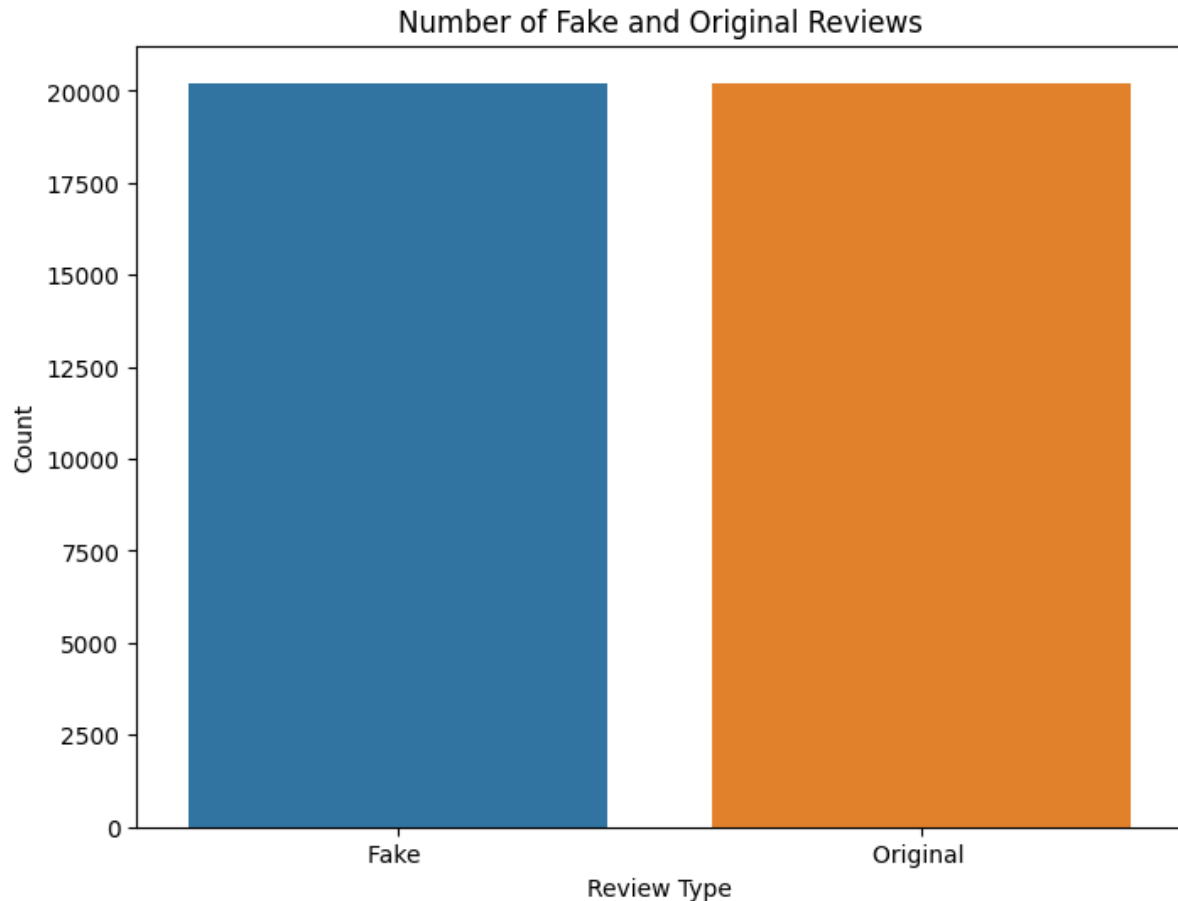


Figure 3 Data Division

3.2.2 Data Pre-processing

Before performing the feature extraction step, the data must undergo several cleanings, such as punctuation removal to remove punctuation marks from the review's text (? , ! , ; , ; , " , . ,) , stop word removal from removing article words from the review sentences ('the,' 'a,' 'an,' 'in,'), eliminating extraneous words and characters from the entire dataset, and data tokenization to separate each sentence of the review contents into separated words, keywords, phrases.

text	target	tokenized	stopwords_removed
text_	0	[]	[]
Love this! Well made, sturdy, and very comfortable. I love it! Very pretty	1	[Love, this, Well, made, sturdy, and, very, comfortable, I, love, it, Very, pretty]	[Love, Well, made, sturdy, comfortable, I, love, Very, pretty]
I love it, a great upgrade from the original. I've had mine for a couple of years	1	[love, it, a, great, upgrade, from, the, original, I, had, mine, for, a, couple, of, years]	[love, great, upgrade, original, I, mine, couple, years]
This pillow saved my back. I love the look and feel of this pillow.	1	[This, pillow, saved, my, back, I, love, the, look, and, feel, of, this, pillow]	[This, pillow, saved, back, I, love, look, feel, pillow]
Missing information on how to use it, but it is a great product for the price! I	1	[Missing, information, on, how, to, use, it, but, it, is, a, great, product, for, the, price, I]	[Missing, information, use, great, product, price, I]

Figure 4 Data Preprocessing

As it can be seen in the figure 3 above, the text feature that tells us about the review the customer gave which is in text form and the target variable which indicates whether the review is fake or real, after that the text review is tokenized in separate words followed by the step of stop words removal in which the stop words that aren't useful and decrease the accuracy of model are removed, examples are this, and, it etc.

3.2.3 Feature Selection:

In our study on the detection of fake reviews in the e-commerce industry, we utilized feature selection to build an accurate, efficient model. Our strategy intended to identify attributes unique to fake reviews in the provided dataset. The feature selection process for detecting fake reviews on e-commerce platforms was comprised of the following criteria:

Length count: As a feature, we counted the review's total length.

Bigram Type: Applying a bigram probability model to determine how likely it was that the next word in the review would also be a bigram was estimated. Without looking at the whole document, we can use the Markov model to guess the probability of the next word.

Sentiment word count: We figured out how people felt about the words in the ratings by comparing them to lists of positive and negative keywords we had already made. This helped us determine if the review said good or bad things.

Noun and verb count: Part-of-speech tagging, like the Natural Language Toolkit (NLTK), was used to find and count the number of nouns and verbs in the reviews. This was part of the process of choosing which features to use.

We split the data into 80% of the training and 20% of the testing samples to see how well our model worked. With this split's help, we could train the model on a sizable chunk of the data and assess its precision with the remaining samples.

3.2.4 Term frequency-inverse document frequency (TF-IDF)

Term frequency-inverse document frequency, or TF-IDF, is used in text classification systems. It is seen as one of the ways to pull out and show features. It is used to find information and natural language processing. The TF-IDF statistical method also ranks how vital a term or word is to each text in the dataset. It has two parts: term frequency, which looks at how often certain words appear in different texts to see how similar they are.

3.2.5 Machine Learning Techniques

To safeguard their reputation and ensure the authenticity of their reviews. Companies are increasingly turning to machine learning methods. These techniques offer both accuracy and speed in identifying fake reviews. Machine learning algorithms can detect such fraudulent activities effectively, making them invaluable for businesses. Various approaches like sentiment analysis, natural language processing (NLP), and analyzing review structure, can be employed to uncover fake reviews. Sentiment analysis involves evaluating user-generated text based on the emotions it evokes in readers. NLP methods delve into the meaning of words or phrases within a review to determine their trustworthiness. Examining factors like review

length, language usage, and punctuation can also help identify genuine reviews. Additionally, machine learning algorithms enable organizations to identify trends in customer feedback that may indicate a fake review. With this insightful information, companies can accurately gauge the likelihood of a review being deceptive and take appropriate action accordingly. Machine learning algorithms are also effective in identifying skewed or factually incorrect reviews. Implementing this technology empowers businesses to pinpoint fake reviews more reliably while protecting their reputation and customers from unscrupulous individuals. By utilizing these machine learning techniques, companies possess a valuable tool set for detecting and combating fake reviews swiftly and accurately—thus fostering trust among customers when they rely on internet reviews for informed purchasing decisions. This part discusses the supervised algorithms used to determine whether the review text is true or false. Before the machine learning classifier's learning, the dataset is split into TF-IDF features, a training set, and a test set. In this project, different types of supervised classifiers will be used: a support vector machine (SVM), a Naive Bayes (NB), a random forest (RF), a logistic regression (LR), etc.

3.2.5.1 Machine Learning Models and Evaluation:

The main goal of this section was to find out how well different machine learning models worked in the e-commerce business at spotting fake reviews. We used a rigorous methodology to achieve this, which included preprocessing the data, choosing useful features, and testing each classifier thoroughly. In the beginning, we picked and set up several classifiers, such as `XGBClassifier`, `CatBoostClassifier`, `LinearSVC`, `AdaBoostClassifier`, `KNeighborsClassifier`, etc., that are known for being good at spotting fake reviews. When each classifier was first set up, specific hyperparameters were carefully picked. We used 5-fold cross-validation, a well-known technique that evaluates a model's generalization capacity, to analyze the performance of these classifiers. The dataset was divided into five equal folds during this phase. We developed a pipeline for each classifier that combined a `TfidfVectorizer` with the specific classifier.

By giving each word a weight depending on its frequency and significance, the `TfidfVectorizer` converted the text input into numerical characteristics. The pipeline was then fitted using the training data, and each fold's test data was used to evaluate it. The performance was evaluated using criteria like `roc_auc`, which assesses the model's capacity to discriminate between positive and negative examples. We obtained several `roc_auc` scores through cross-validation, which enabled us to compute the mean `roc_auc` score across the folds and derive an overall performance metric. The `roc_auc` scores' standard deviation also shed light on the consistency of the model's performance. Based on the findings, we ordered the classifiers according to their `roc_auc` scores in descending order. As a result, we determined which classifiers were most effective in spotting fake reviews. The rankings guided choosing the best classifier for the e-commerce dataset's fake review identification. We also carried out individual assessments of the two best classifiers, `XGBClassifier` and `LinearSVC`. The `TfidfVectorizer` and the appropriate classifier were combined to generate bundled pipelines for each classifier. These pipelines were tested using different test data after being tested on the training data. We determined the number of evaluation metrics, including accuracy, precision, recall, and

roc_auc scores, by contrasting the projected labels with the actual labels. These metrics thoroughly evaluated how well each classifier performed in distinguishing between fake and genuine reviews. According to the evaluation results, the XGBClassifier and LinearSVC models displayed high accuracy, precision, recall, and roc_auc scores. This demonstrated their ability to successfully differentiate between fake and real reviews in the e-commerce dataset. To summarize, our work used a strong technique, including data preparation, informative feature selection, and thorough evaluation of multiple classifiers. The top-performing models were XGBClassifier and LinearSVC, proving they could spot fake reviews accurately. By utilizing these models, e-commerce companies can improve user experience by increasing trust and dependability by automatically identifying fake reviews. Our study's conclusions advance the field of fake review detection and offer helpful information for e-commerce applications.

3.3 Classification Models

3.3.1 SVM

A supervised machine learning algorithm is called SVM. The algorithm is highly good at classifying data since it is built to maximize the margin between the two classes. SVM is useful for sentimental and textual analysis since it can consider the context of words to grasp their meaning better. By examining the review language and the rating, the SVM can assist us in identifying fake reviews. It's probably a fake review if it's positive but gets a low rating. A potent supervised learning system called Support Vector Machine (SVM) identifies and categorizes fake reviews. The Support Vector Machine (SVM) is a highly effective machine learning technique that utilizes input data to determine if a review is fake using preset characteristics. The SVM can map and classify data points. Ultimately creating an ideal hyperplane that separates the two classes of data points. The algorithm takes into account any misidentification of points. Resulting in a precise and reliable conclusion. Additionally, SVM is applied in various machine-learning tasks such as handwriting recognition, picture identification, and text classification. Therefore, Due to its accuracy and efficiency. SVM has proven to be a valuable method for detecting fake reviews. The Linear Support Vector Classifier (LinearSVC) is based on the concept of Support Vector Machines (SVM). This powerful machine-learning approach searches for the best hyperplane within the feature space to separate multiple classes. The "linear" aspect of LinearSVC refers to its assumption that the data can be separated linearly or by using a straight line or hyperplane in higher dimensions. Consequently, LinearSVC is particularly well suited for high-dimensional text data often encountered in various applications. The LinearSVC can be trained to detect fake reviews using text-based features extracted from the reviews. To determine the relative importance of each word in the review, it might employ TF-IDF values or bag-of-words representations. The LinearSVC can efficiently discern between genuine and fraudulent reviews by learning a linear decision boundary. It's important to remember, however, that text data, which can be highly delicate and complicated, may only occasionally abide by the premise of linear separability. Therefore, performing exploratory data analysis and correctly preprocessing the

data before training the model is crucial. This could involve techniques like text cleaning, vectorization, and tokenization.

3.3.2 Extreme Gradient Boosting:

The XGBClassifier uses the Extreme Gradient Boosting technique, also called the XGBoost algorithm. As a component of an efficient ensemble learning technique, this algorithm generates predictions using a variety of decision trees. This method's "boosting" capability refers to its ability to combine numerous weak learners, frequently decision trees, to create a robust, trustworthy model. XGB classifier has the ability to effectively manage large data sets due to its efficient processing and memory utilization. Additionally, it can automatically handle the presence of missing values, which is a common issue encountered in real-world data. In the context of detecting fake reviews, the XGBClassifier can be employed to decipher complex data patterns. This classifier can identify linguistic trends in reviews, including repeating specific terms or phrases frequently used in fake reviews. Furthermore, it can consider other factors, such as the reviewer's background, the timing of the review, and the product details. By combining these features, the XGBClassifier can create a robust model that successfully distinguishes between real and fake reviews. The model's performance can be improved by changing hyperparameters like the learning rate, the maximum depth of the trees, and the number of trees in the ensemble. The model's forecast accuracy can be increased by using this fine-tuning approach to help it better understand the subtleties in the data.

3.3.3 CatBoost Classifier:

The CatBoostClassifier is a gradient-boosting technique that is specially made to handle categorical features well. Features classified as categorical are those that only have a small range of possible values. The product category, the review platform, or the reviewer's location are examples of categorical features that can be used to identify fake reviews. These qualities can offer important information regarding the propensity for false reviews. For example, some product categories might be more vulnerable to fraudulent reviews, or some platforms might have stricter review criteria that discourage them. The CatBoostClassifier is a practical option for this assignment because it can handle these features without the need for extensive feature engineering. The CatBoostClassifier can also handle text features by transforming them into numerical representations. For instance, it can transform the review text into a format that can be utilized for training using strategies like bag-of-words or word embeddings. These methods can extract the text's semantic content and transform it into a format that the algorithm can comprehend. The CatBoostClassifier can deliver reliable results by accurately capturing the connections between these traits and the target variable.

3.3.4 Adaboost Classifier:

The powerful ensemble machine learning technique, AdaBoost, or adaptive boosting, combines several weak classifiers to create a robust classifier. Simply put, a weak classifier is one that performs poorly but is still better at guessing at random. AdaBoost is adaptive in that it modifies succeeding weak learners in favor of instances that prior classifiers incorrectly classified. AdaBoost has the potential to be very useful in the situation of identifying fake reviews. A well-known example of a binary classification problem is the challenge of fake review identification, where reviews are categorized as "genuine" or "fake." Reviews

frequently contain high-dimensional and noisy text data, making it difficult for a single classifier to perform well. In this regard, AdaBoost excels. AdaBoost first fits a base classifier (such as a decision tree) on the original dataset to fit additional classifiers on the same dataset. However, the weights of mistakenly categorized instances are changed such that subsequent classifiers concentrate more on challenging cases. This cyclical procedure continues until a predetermined number of poor learners are produced or until no more progress can be accomplished. The weighted average of the weak classifiers makes up the final model. There are many benefits of utilizing AdaBoost to identify fake reviews. First off, overfitting is less likely with AdaBoost. Even though it combines multiple weak learners to create a complicated model, it typically does not overfit the training set of data. Second, a robust classifier integrates various straightforward models to provide a stable, balanced prediction model. Since reviews often contain high-dimensional, noisy text data, it can effectively handle this. AdaBoost is also quick, adaptable, and simple to program. It is an excellent option for non-expert users because it needs minimal adjusting of parameters or settings. It can be integrated with any learning technique and utilized with various data formats, including text, numeric, and discrete data. AdaBoost, a powerful and adaptable machine learning algorithm, can successfully address the issue of false review identification. It is an effective weapon in the fight against fraudulent reviews because of its capacity to concentrate on cases that are more difficult to classify and merge several weak classifiers into one strong one.

3.3.5 Logistic Regression

The sophisticated logistic regression (LR) statistical technique for binary classification makes it the perfect instrument for identifying authentic reviews from fake ones. Unlike linear regression, which presumes a linear relationship between the dependent and independent variables, logistic regression (LR) models the likelihood that a given input point belongs to a specific category using a logistic function, often known as the sigmoid function. The two categories in the context of fake review detection would be "real" and "fake." The logistic function is appropriate for understanding the output as a probability since it maps any real-valued number onto the (0, 1) interval. The program generates a linear equation from the observed data and predicts the log odds of the output is '1' (in this case, a fake review) rather than the direct result. The logistic regression model's weights or coefficients describe the change in log odds brought on by the relevant feature. Maximum likelihood estimate is a technique used to learn these weights from the data. Once trained, the model may take a fresh review, apply the learned weights to the review's attributes, and then pass the outcome through the logistic function to obtain a probability. The review is labeled as fake if this likelihood exceeds a predetermined cutoff point, which is often 0.5. LR is beneficial when the decision boundary between the classes is roughly linear. Additionally, it helps when you require simple models to comprehend and apply or when you need probabilistic findings.

3.3.6 Naive Bayes

Using Bayes' theorem with strong (naive) independence assumptions between the features, Naive Bayes (NB) is a classification technique. It is a straightforward yet effective technique for classification and predictive modeling, and it performs well when working with massive

datasets, which makes it a suitable choice for fake review identification. The Naive Bayes classifier determines the likelihood that a review is fake given its characteristics (words or phrases in the review) in the context of fake review detection. It achieves this by using Bayes' theorem, which explains the link between conditional probabilities of statistical quantities. The 'naive' assumption in Naive Bayes is that each pair of features is independent. Although it is a strong assumption, it simplifies computation, and Naive Bayes classifiers frequently outperform other algorithms in real-world applications. The technique is robust to irrelevant features common in text data and can handle both continuous and discrete data. Additionally, it is computationally effective, making high-dimensional datasets ideal for it. Naive Bayes can be an effective tool for fake review identification when used in conjunction with methods like bag-of-words or TF-IDF for text feature extraction..

3.3.7 Decision Tree

A non-parametric supervised learning technique for classification and regression is called a decision tree (DT). Because they more closely resemble human decision-making than other methodologies, they are intuitive and simple to understand. Decision trees can be used to learn a set of questions to discriminate between authentic and fake reviews in the context of fake review detection. The internal nodes of the tree represent questions regarding the input (such as, "Does the review contain the word 'unbelievable'?"), and the branches from the node represent potential solutions. Each leaf node is a decision, either "real" or "fake," in this example. The tree is designed in a way that attempts to reach the most accurate conclusions with the fewest number of queries. Decision trees' interpretability is one of their key benefits. In applications like fake review identification, where it would be necessary to justify the choice to a human, it can be helpful if the model's predictions are simple to grasp. But when the tree is exceedingly deep, decision trees might be vulnerable to overfitting.

3.3.8 Random Forest

Random Forest (RF) is an ensemble learning technique that creates a large number of decision trees during training and outputs the class that represents the mode of the classes of the individual trees. It's a potent method that frequently yields better outcomes than using just one decision tree. A random forest model would train a number of decision trees on various subsets of the training data in the context of fake review detection. The class of a review is decided by a vote from each tree in the forest. The result of the random forest is the class (actual or fake) with the most votes. Random forests have a number of benefits. They don't require feature scaling and can handle a lot of features. They provide a more accurate forecast than a single decision tree as they minimize variance by averaging the results of several trees. As a result, they are less likely to overfit the training set. They may be more challenging to understand than single decision trees because the forecast depends on the votes of numerous trees.

3.3.9 K-Nearest Neighbours

One of the many uses for the adaptable and well-known K-Nearest Neighbors (KNN) algorithm in machine learning is detecting fake reviews. The usage of features to group related instances is its main idea. KNN can be used to identify patterns in fake reviews and classify them as authentic or fake based on the characteristics of their near neighbors. The approach calculates the distance between a particular review and its neighboring reviews in the feature

space when using KNN to identify fake reviews. This distance is typically computed using a distance function, such as the Euclidean distance or the cosine similarity. By considering the classification of its nearest neighbors, KNN assigns a label to the target review and classifies it as either genuine or fake. The simplicity and effectiveness of KNN are its main benefits, which makes it ideal for identifying patterns in large datasets, including linguistic elements typically observed in reviews. By comparing and contrasting the closest neighbors' traits, KNN can accurately distinguish between genuine and fraudulent reviews. Because of its low processing cost, KNN is also useful for real-time or online applications requiring quick decisions. KNN can be pretty helpful in fake review detection for spotting suspicious reviews. You can create a system that successfully distinguishes between real and false reviews by utilizing KNN's capacity to spot patterns and similarities. KNN can offer helpful insights for spotting potential instances of fake reviews by considering features, including review content, reviewer information, and product specifics. It can be combined with other algorithms in a comprehensive framework to improve the system's overall performance and accuracy.

3.3.10 Stochastic Gradient Descent (SGD) Classifier:

The stochastic gradient descent, also known as SGD, is a method that may be used to learn linear classifiers that have convex loss functions, such as Support Vector Machines and Logistic Regression. This method is both powerful and efficient. It has been demonstrated to effectively address large-scale and dense machine learning challenges, frequently experienced in classifying texts and natural language processing. An objective function with smoothness qualities can be optimized using SGD, an iterative approach. It is a stochastic approximation of the gradient descent algorithm, in which the actual gradient computed from the complete dataset is substituted with an estimate based on a portion of the data chosen at random. This strategy decreases the amount of calculation needed, enabling faster iterations at the expense of a slightly slower convergence rate. The SGD classifier is a linear classifier trained using the SGD algorithm. Some examples of linear classifiers include SVM and logistic regression. It does this by putting regularized linear models that learn using stochastic gradient descent. Each sample's loss gradient is estimated, and then the model is updated in an iterative manner using a decreasing learning rate. Through its use of the partial fit approach, SGD also offers minibatch learning, which can take place online or outside of the core. The use of stochastic gradient descent comes with several benefits, including the same piece of code can be utilized to train multiple different machine learning models thanks to SGD's efficiency. It is possible to start making progress immediately by computing the gradient using one data point at a time, and it will continue to improve as additional data becomes available. SGD is simple to implement and provides flexibility for code optimization, opening the door to chances for customization and fine-tuning. SGD's ease of use also makes it simple to use.

Because it is both efficient and effective, SGD can be useful in situations involving detecting fake reviews. Text data is frequently high-dimensional, and SGD is able to handle such data in an efficient manner, hence giving a model that is capable of identifying reviews as authentic or fake. Because of its linear character, SGD is ideally suited for problems involving text classification, in which the decision boundary is typically linear. In addition, because SGD can manage enormous amounts of data, it is an excellent choice for situations requiring a significant number of reviews to be completed. To improve model performance, SGD may be integrated with several other methods, such as Multilayer Perceptron (MLP), in a hybrid

modeling approach. It is possible that using SGD to optimize the weights of the MLP results in a model that is superior in its ability to identify fake reviews, as proposed in this research.

3.4 Neural Network Models:

3.4.1 Multilayer Perceptron(MLP):

A sort of feedforward artificial neural network model that maps sets of input data onto sets of suitable outputs is the multilayer perceptron (MLP). Each layer of nodes in an MLP is completely connected to the layer below it in a directed graph. The nodes, often called neurons, employ a nonlinear activation function that establishes a neuron's output in response to an input or combination of inputs. An MLP can be used to categorize reviews as being real or fake. The TfidfVectorizer is the initial stage in the workflow. This method for turning text data into a useful numerical representation is utilized in text mining. Inverse Document Frequency rates the words that often recur across all documents, while Term Frequency measures how frequently a term appears. Together, these scores draw attention to words that are more intriguing and common in one document than another. To determine whether a review is fake, the MLP classifier learns to link these numerical representations with the labels. The MLP achieves this through a technique known as backpropagation, which modifies the weights of the connections in accordance with the error at the output. Until the error is minimized and the network can precisely predict the class of the incoming data, the MLP keeps adjusting these weights. The benefit of employing an MLP for this task is that it can learn nonlinear models and can use `partial_fit` to learn models online in real time. Because of its versatility and ability to handle complicated patterns in data, the MLP is a practical algorithm for managing text data like reviews.

3.4.2 Convolutional Neural Network(CNN):

Convolutional Neural Networks (CNNs) are a subset of deep learning models that are applied mainly to processing grid-like data, such as photographs. Their use has been expanded to include text and other forms of data, though. A multilayer perceptron variant used by CNNs is made to require less preprocessing. You utilize a CNN at work to handle reviews. The embedding layer, which is the first layer in model, turns the words in reviews into dense vectors of a set size. This method translates the semantic ties between words into geometric relationships in a high-dimensional space, where words are represented as vectors. The data is processed using a convolutional layer that follows using filters and a sliding window. The purpose of this layer is to automatically and adaptively learn feature spatial hierarchies. When applied to text data, these can be viewed as learning the keywords or phrases that signify a fake review. The GlobalMaxPooling1D layer preserves the most significant feature (the one with the most outstanding value) for each feature map by converting the output of the convolutional layers to a single vector. The final prediction is then produced using this vector as input into a dense layer. Using a CNN for this task has the benefit for collecting local and global patterns in the input data, which, in the case of text data, might equate to the comprehension of phrases and sentences. CNNs are, therefore, extremely useful for this task.

3.4.3 Gated Recurrent Unit(GRU):

Recurrent neural networks (RNNs) with gated recurrent units (GRUs) have gating units comparable to those in Long Short-Term Memory (LSTM) cells but with a more straightforward topology. GRUs are made to capture dependencies on various time scales in sequential data in a flexible manner. The embedding layer, which is the first layer in your model, turns the words in your reviews into a dense vector of a set size. This method translates the semantic ties between words into geometric relationships in a high-dimensional space, where words are represented as vectors. The GRU layer subsequently processes these vectors. A GRU has gating units that regulate the flow of information inside the unit without relying on the memory cells of the LSTM, in contrast to conventional recurrent neural networks. This makes it effective at capturing the temporal dependencies in the text data by allowing it to maintain relevant information in the input data and eliminate the irrelevant ones. To output the likelihood that a review is fake, the output of the GRU layer is then sent through a dense layer with a sigmoid activation function. Using a GRU for this purpose has the advantage of being able to capture the temporal relationships in the text data, which can correspond to the comprehension of the context and semantic meaning of phrases. Due to this, GRUs are very useful for this task.

3.5 The Ensemble/Hybrid Model:

When two of above classifiers(Deep learning model and Machine learning) are combined into a singular model using a VotingClassifier, the power of ensemble learning becomes evident. In this case we made an ensemble of SGD classifier from machine learning and MLP from deep learning. Ensemble learning methods, such as the VotingClassifier, function by combining the predictions of multiple models. The underlying principle is that a group of weak learners can constitute a single strong learner. In this instance, the VotingClassifier was configured to use 'soft voting,' which predicts the class label using the argmax of the sums of the predicted probabilities instead of a straightforward majority vote ('hard' voting).

Cross-validation was applied to assess the ensemble model further. Cross-validation is reliable for evaluating a model's performance on unobserved data. It functions by dividing the dataset into 'k' subsets and training the model 'k' times, for example, in this case, k=5, using a different subset as the test set and the remaining data as the training set each time. The cross-validation results are then averaged to determine the overall efficacy of the model.

3.6 Chapter Summary:

This section will look at the study's methodology, including the instruments and measures employed to spot fraudulent reviews (such as total reviews, positive reviews, reviewer deviation, review length, and the LIWC instrument). Next, the data pipeline is discussed, including data collection, pre-processing, feature selection, and machine learning techniques like the term frequency-inverse document frequency (TF-IDF) method. Support Vector Machines (SVM), Extreme Gradient Boosting (XGB), CatBoost Classifier (CB), Adaboost Classifier (Adaboost), Logistic Regression (LR), Naive Bayes (NB), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), and Stochastic Gradient Descent (SGD) Classifier are only some of the classification models investigated. Multilayer Perceptron

(MLP), Convolutional Neural Network (CNN), and Gated Recurrent Unit (GRU) are the neural network models covered in this chapter's last section. The research instruments and measures, data pipeline, classification models, and neural network models used in the study are all detailed in this chapter. Also the ensemble model was described and how that worked is discussed in this chapter.

ANALYZE2GETHER.COM

Chapter 4

Results and Analysis:

It is essential to measure the accuracy of all machine learning techniques when using them for detecting fake reviews. This is because if a machine learning model is not accurate, then any predictions or classifications that it makes about a review will be wrong or misleading. For this reason, it is extremely important to ensure that a model's accuracy is properly assessed before it is used in a production environment. To do this, data scientists should use various evaluation metrics such as precision and recall evaluating the performance of a model so that they can have confidence in its accuracy. To guarantee that the model keeps producing accurate results, it is also crucial to retrain and validate the model occasionally using fresh samples from older or new datasets. By doing this, we can be confident that the machine-learning models are constantly updated with the latest facts and trends.

4.1 Machine Learning Classifiers:

Table 5 Machine Learning Classifiers

Classifier	Accuracy	Precision	Recall	ROC/AUC Score
LinearSVC	87.44%	87.76%	87.67%	87.43%
RidgeClassifier	87.44%	88.35%	86.90%	87.45%
RandomForestClassifier	86.45%	84.88%	89.44%	86.38%
LGBMClassifier	86.00%	86.96%	85.40%	86.02%
SGDClassifier	85.83%	88.81%	82.72%	85.90%
XGBClassifier	85.56%	87.61%	83.59%	85.60%
MultinomialNB	85.25%	83.05%	89.41%	85.15%
AdaBoostClassifier	78.20%	80.30%	76.05%	78.25%
DecisionTreeClassifier	77.43%	77.94%	77.96%	77.42%
BernoulliNB	74.37%	67.89%	94.68%	73.89%
KNeighborsClassifier	64.07%	59.38%	94.17%	63.35%
ExtraTreeClassifier	69.99%	69.70%	73.12%	69.91%

The SGDClassifier correctly classified 85.83% of the reviews, according to its accuracy score of 85.78%. The high precision of 88.81% also indicated a low rate of false positives that it reached. The recall of 82.72% indicates that the system correctly identified 82.72% of the genuine positive reviews. The model did a fantastic job differentiating between positive and negative reviews, earning an 85.90% ROC/AUC score. It was possible to achieve the accuracy of the XGBClassifier, which was 85.56%, slightly less accurate than the SGDClassifier. With a precision of 87.61%, there was a low rate of false positives. The recall of 83.59% indicates that the model correctly identified a significant portion of the positive reviews. The classification of reviews performed well overall, as seen by the ROC/AUC score of 85.60%. With a score of 87.44%, the LinearSVC model outperformed all other classifiers in terms of accuracy. Additionally, it has a high precision of 87.76%, indicating a low rate of false positives. The recall of 87.67% means that the model correctly detected a significant fraction of the positive reviews. Excellent performance in differentiating between positive and negative

reviews is indicated by the ROC/AUC score of 87.43%. With an accuracy of 85.25%, the MultinomialNB classifier fell just short of the prior models. Compared to the earlier models, it had a precision of 83.05%, indicating a substantially greater false positive rate. However, it managed to recall 89.41% of the positive reviews, meaning that most of them were accurately identified. The ROC/AUC score of 85.15% indicates an excellent overall performance.

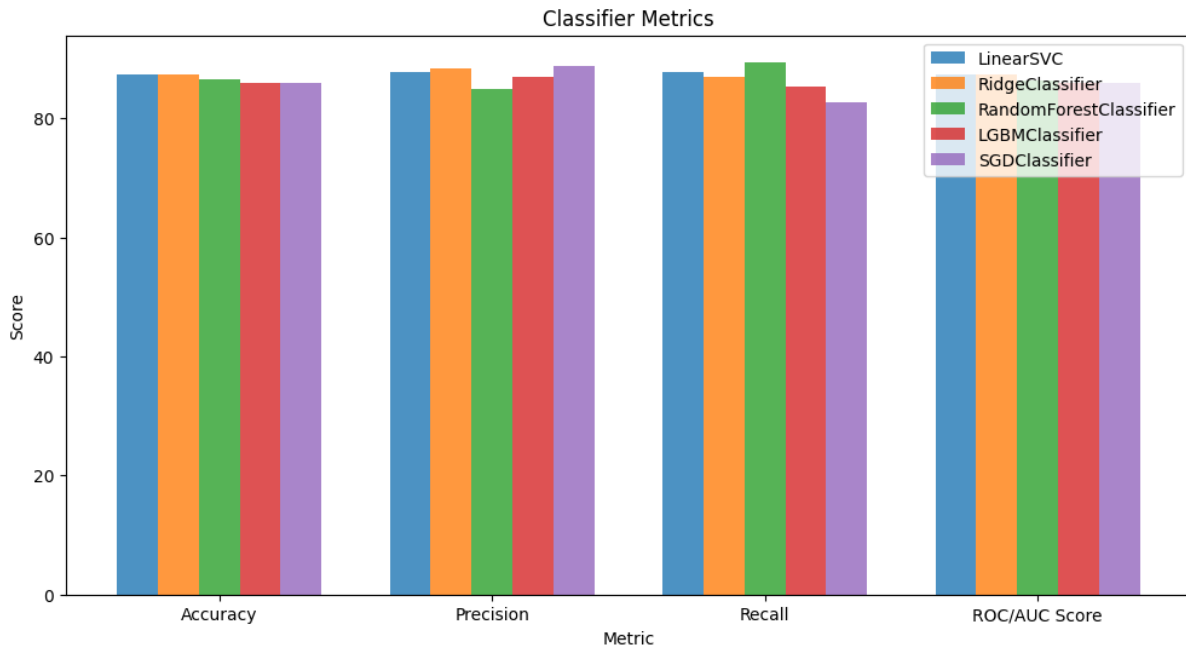


Figure 5 Machine Learning Algorithms Results

With an accuracy of 86.00%, the LGBMClassifier outperformed the MultinomialNB classifier. It had a precision of 86.96%, indicating a low false positive rate. The recall of 85.4 % indicates that the model properly detected a sizable percentage of the positive reviews. The ROC/AUC score of 86.02% indicates an excellent overall performance. The accuracy of the RandomForestClassifier, which was marginally better than the LGBMClassifier, was 86.45%. Compared to the earlier models, it had a precision of 84.88%, indicating a substantially greater false positive rate. However, the fact that it was able to recall 89.44% of the positive reviews means that a significant number of them were accurately identified. The ROC/AUC score of 86.38% shows an excellent overall performance. The accuracy of the DecisionTreeClassifier, which was lower than that of the earlier models, was 77.43%. Compared to the earlier models, it had a precision of 77.94%, indicating a substantially greater false positive rate. The recall of 77.96% indicates that the model properly detected a sizable percentage of the positive reviews. An average performance is shown by the ROC/AUC score of 77.42%. The ExtraTreeClassifier has the lowest accuracy of all the classifiers, scoring 69.99%. Compared to the earlier models, it had a precision of 69.70%, indicating a substantially greater false positive rate. The recall of 73.12% indicates that the model correctly detected a sizable percentage of the positive reviews. The performance is below average, as indicated by the ROC/AUC score of 69.91%. The AdaBoostClassifier's accuracy of 78.20% was lower than that of most earlier models. The precision was 80.30%, indicating a low false positive rate. The recall of 76.05% indicates that the model properly detected a sizable percentage of the positive reviews. An average performance is shown by the ROC/AUC score of 78.25%. With a score of 64.07%, the KNeighborsClassifier has the lowest accuracy of any classifier.

Compared to the earlier models, it had a precision of 59.38%, indicating a substantially greater false positive rate. However, it managed to acquire a high recall of 94.17%, indicating that it accurately identified a significant number of the positive reviews. The performance is below average, as can be noticed by the ROC/AUC score of 63.35%. With a score of 87.44%, the RidgeClassifier equaled the LinearSVC model's accuracy. It had a precision of 88.35%, indicating a low false positive rate. The recall of 86.90% means that the model accurately detected a significant majority of the positive reviews. Excellent performance in differentiating between positive and negative reviews is indicated by the ROC/AUC score of 87.45%. The accuracy of the BernoulliNB classifier was 74.37%, which is less than most of the earlier models. Compared to the earlier models, it had a precision of 67.89%, indicating a substantially greater false positive rate. However, it managed to acquire a high recall of 94.68%, suggesting that it accurately identified a significant number of the positive reviews. A below-average performance is indicated by the ROC/AUC score of 73.89%.

4.2 Deep Learning Algorithms:

Table 6 Neural Network Results

Network	Accuracy	Precision	Recall	AUC/ROC
MLP	85.66%	84.74%	87.77%	85.61%
CNN	91.47%	90.47%	93.14%	91.43%
GRU	91.11%	89.18%	94.03%	91.04%

Whereas, The analysis of the performance and suitability of three deep neural networks, namely Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), and Gated Recurrent Unit (GRU), for detecting fake reviews reveals fascinating insights.

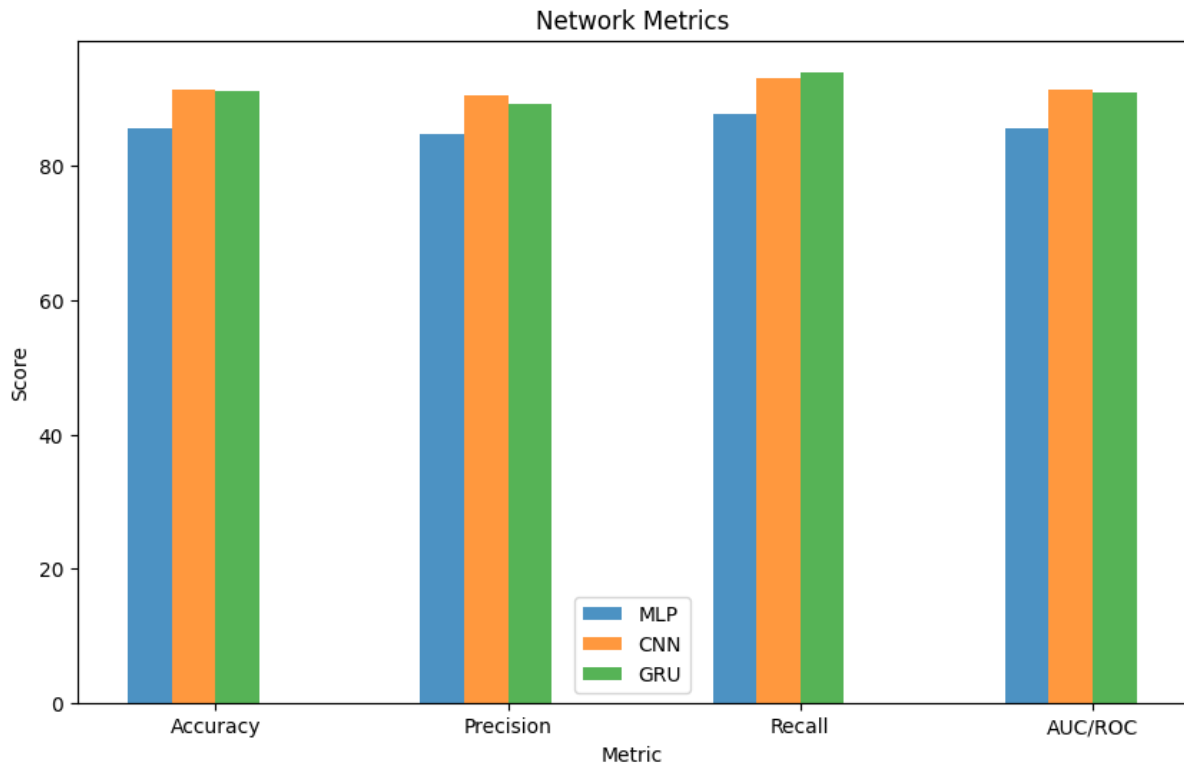


Figure 6 Neural Network Results

Referred to the figure 5 the MLP obtained an accuracy of 85.66%, indicating that it correctly identified 85.66 %of reviews as fake or genuine. The precision of 84.74% indicates that 84.74% of the reviews categorized as fake were fake. The recall rate of 87.77% demonstrates that the MLP accurately identified 87.77% of the fake reviews. The AUC/ROC score of 85.61% reflects its capacity to differentiate between fake and authentic reviews. The CNN outperformed the MLP with an accuracy of 91.47%. This suggests that CNN's predictions were more accurate overall. The precision of 90.47% indicates a reduced false positive rate, i.e., a smaller proportion of authentic reviews were misclassified as fake. The recall rate of 93.14% indicates that CNN had a reduced false negative rate and correctly identified a more significant proportion of the fake reviews. The AUC/ROC score of 91.43% demonstrates the capacity of this classifier to rank the reviews, in consideration of its effectiveness. The GRU did similarly well, achieving an accuracy of 91.11%. The 89.18% precision reports a low false positive rate, while the recall of 94.03% indicates a low false negative rate. The GRU's ability to distinguish between fake and real reviews was further validated by achieving an AUC/ROC score of 91.04 percent.

The results showcased herein highlight the effectiveness of deep neural networks in uncovering fake reviews. Deep neural networks are highly suitable for this purpose thanks to their capability to identify intricate patterns and connections within textual data. Moreover, they understand the non-linear interactions necessary to distinguish between authentic and fake reviews exhibiting intricate patterns. Additionally, CNNs automatically extract meaningful characteristics from text data using convolutional layers, demonstrating the effectiveness of deep neural networks in feature extraction. This makes it possible for them to identify significant patterns and structures that signify the reliability of reviews.

Additionally, the GRU can identify temporal connections in sequential data because it is a recurrent neural network. This is particularly helpful for examining reviews' sequential nature and seeing trends that show whether the information is fake or legitimate. Additionally, deep neural networks display great generalization to unseen data, enabling them to make precise predictions on new reviews that weren't part of the training set. This is crucial for real-world applications where recent fake reviews could manifest. Deep neural networks can effectively handle enormous volumes of data and are scalable. This is essential for processing several reviews, which happens frequently in real-time applications.

MLP, CNN, and GRU findings demonstrate that deep neural networks can successfully identify fake reviews. They are valuable tools for addressing the challenges of recognizing misleading information on online platforms because of their capacity to catch intricate patterns, extract important aspects, model temporal connections, generalize to unforeseen data, and effectively handle vast amounts of data.

4.3 The Ensemble Model:

The MLP (Multilayer Perceptron) and SGDClassifier (Stochastic Gradient Descent Classifier) classifiers excelled in their respective odds. The MLP accurately classified 85.66% of the cases, according to its accuracy score of 85.66%. The accuracy of the SGDClassifier, in comparison, was marginally higher at 85.73%. These high accuracy values show that both models could successfully learn from the training data and accurately categorize a sizable number of examples.

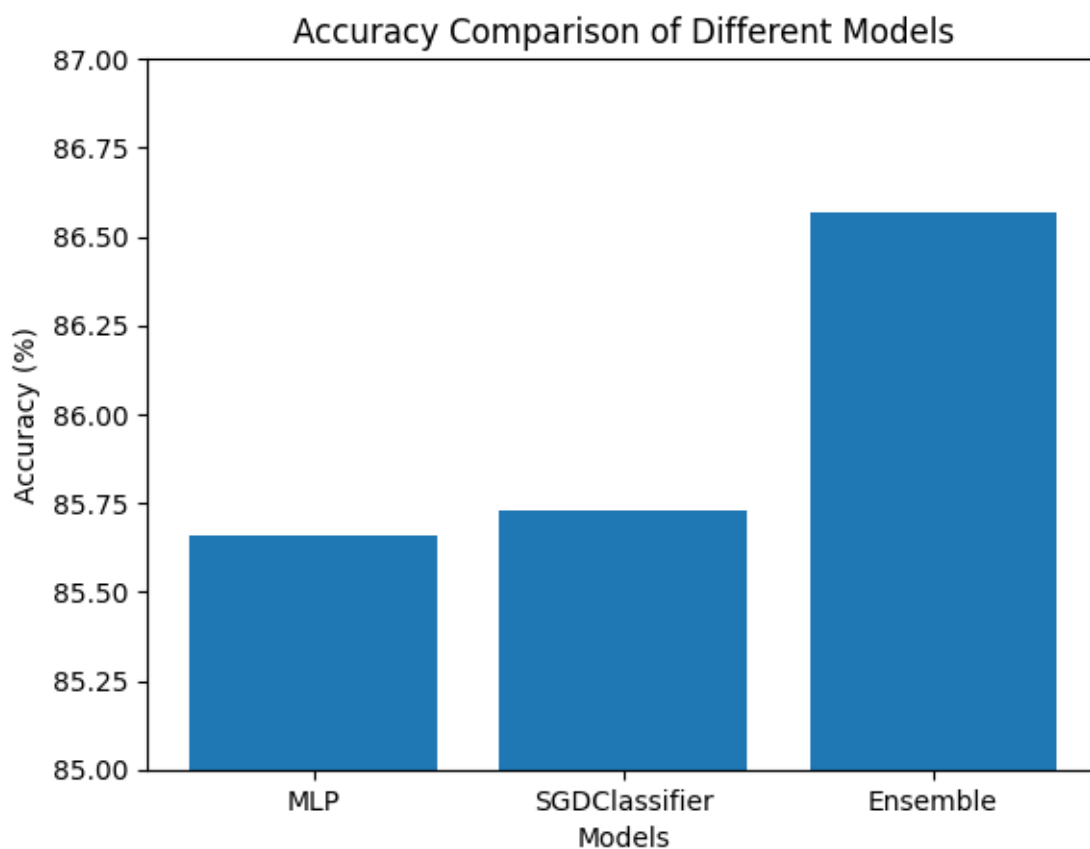


Figure 7 Ensemble Model Metrics

The ensemble model's accuracy increased to 86.57% , a significant increase over the accuracy of the individual classifiers. This demonstrates how ensemble methods improve model performance. Using the strengths of multiple models, the ensemble model can accomplish greater accuracy and generalization than any individual model.

The ensemble model's mean cross-validation score was 85.56%, which was marginally lower than the model's accuracy. This discrepancy may result from data variation or indicate that the ensemble model is marginally overfitting the training data. When a model learns the training data so well that it performs unfavorably on unseen data, this is known as overfitting. However, the reasonably slight difference between the accuracy and the mean cross-validation score indicates that if overfitting is present, it is not severe.

Hypothesis H1, which stated that the machine learning models better detect the fake reviews is rejected as we can see in the findings of our study. The evaluation and analysis did not support the hypothesis that machine learning can better detect the fake reviews.

Hypothesis H2, which stated that the machine learning models better detect the fake reviews is rejected as we can see in the findings of our study. The evaluation and analysis did not support the hypothesis that deep learning can better detect the fake reviews.

Where as **Hypothesis H3**, shows that it is the best fit model for the detection of fake reviews and is supported by the evidence, so it is accepted. The results of the study showed that the proposed hybrid model outperformed both the machine learning and the deep learning models and produced better accuracy then both of the parent algorithms.

This analysis concludes by demonstrating the effectiveness of ensemble methods in machine learning. An ensemble model can accomplish greater accuracy and generalization than any single model by combining the predictions of multiple models. This makes ensemble methods a valuable tool for duties requiring high accuracy.

Chapter 5

Conclusion and Future work:

Currently, no study fully addresses the complex problem of identifying fake reviews. This study focused on an ensemble model for fake review identification to suggest an innovative solution. This study's research has shed important light on how ensemble methods are used in machine learning, particularly when it comes to detecting fake reviews. MLP and SGDClassifier, the two separate classifiers, each displayed strong performance with accuracy values of 85.66% and 85.83%, respectively. The ensemble model, which combined these classifiers' predictions, outperformed the individual models and had an accuracy of 86.57%. This increase in accuracy underlines how effective ensemble approaches are in enhancing the functionality of machine learning models. Cross-validation, a reliable method for gauging a model's performance on unobserved data, was used to test the performance of the ensemble model. The model's accuracy was marginally exceeded by the mean cross-validation score, indicating minimal overfitting. However, due to the comparatively tiny difference between these scores, the ensemble model has strong generalizability and can be anticipated to perform well on untested data.

This work provides various directions for future study. The VotingClassifier's performance may be improved by investigating other ensemble techniques like Bagging, Boosting, and Stacking, and the ensemble model's performance may also be enhanced by adjusting the hyperparameters of the individual classifiers.

Acknowledgement

I am grateful to the National of Computer & Emerging Sciences for providing me with the opportunity to conduct research. I want to thank my supervisors, Dr. Muhammad Bilal Saeed and Mr. Hammad Majeed, for their invaluable assistance, support, and encouragement throughout my research journey.

In addition, I would like to acknowledge Salminen et al. for providing the dataset used in my experiments. I greatly appreciate their decision to make the dataset available to the public for the advancement of research in the identification of fake reviews, as it has significantly contributed to the success of my research.

ANALYZE2GETHER.COM

References

- [1] Abdulqader, M., Namoun, A., & Alsaawy, Y. (2022). Fake online reviews: A unified detection model using deception theories. *IEEE Access*, 10, 128622–128655. <https://doi.org/10.1109/access.2022.3227631>
- [2] Ahmed, H., Traore, I., & Saad, S. (2017). Detecting opinion spam and fake news using text classification. *Security and Privacy*, 1(1). doi:10.1002/spy2.9
- [3] Baishya, D., Deka, J. J., Dey, G., & Singh, P. K. (2021). Safer: Sentiment analysis-based fake review detection in e-commerce using Deep Learning. *SN Computer Science*, 2(6). <https://doi.org/10.1007/s42979-021-00918-9>
- [4] Banerjee, S., Chua, A. Y., & Kim, J. (2015). Using supervised learning to classify authentic and fake online reviews. *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*. doi:10.1145/2701126.2701130
- [5] Bansal, B., & Srivastava, S. (2018). Sentiment classification of online consumer reviews using word vector representations. *Procedia Computer Science*, 132, 1147-1153. doi:10.1016/j.procs.2018.05.029
- [6] Barbado, R., Araque, O., & Iglesias, C. A. (2019). A framework for fake review detection in online consumer electronics retailers. *Information Processing & Management*, 56(4), 1234-1244. doi:10.1016/j.ipm.2019.03.002
- [7] Bhatti, A., Akram, H., Basit, H. M., Khan, A. U., Raza Naqvi, S. M., & Bilal, M. (2020). E-commerce trends during the COVID-19 pandemic. Retrieved November 24, 2022, from https://www.researchgate.net/profile/Ahmed-Khan-67/publication/342736799_E-commerce_trends_during_COVID-19_Pandemic/links/5f04603c458515505091c291/E-commerce-trends-during-COVID-19-Pandemic.pdf
- [8] Cardoso, E. F., Silva, R. M., & Almeida, T. A. (2018). Towards automatic filtering of fake reviews. *Neurocomputing*, 309, 106–116. <https://doi.org/10.1016/j.neucom.2018.04.074>
- [9] Chen, T., Samaranayake, P., Cen, X. Y., Qi, M., & Lan, Y.-C. (2022). The impact of online reviews on consumers' purchasing decisions: Evidence from an eye-tracking study. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.865702>
- [10] Crawford, M., Khoshgoftaar, T. M., Prusa, J. D., Richter, A. N., & Al Najada, H. (2015). Survey of review spam detection using machine learning techniques. *Journal of Big Data*, 2(1). doi:10.1186/s40537-015-0029-9
- [11] Danilchenko, K., Segal, M., & Vilenchik, D. (2022). Opinion spam detection: A new approach using machine learning and network-based algorithms. *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 125–134. <https://doi.org/10.1609/icwsm.v16i1.19278>
- [12] De Maeyer, P. (2012). Impact of online consumer reviews on sales and Price Strategies: A review and directions for future research. *Journal of Product & Brand Management*, 21(2), 132–139. <https://doi.org/10.1108/10610421211215599>
- [13] Ecommerce: The history and future of online shopping (no date) BigCommerce. Available at: <https://www.bigcommerce.com/articles/ecommerce/> (Accessed: November 25, 2022).

- [14]Elmoghy, A. M., Tariq, U., Mohammed, A., & Ibrahim, A. (2021). Fake reviews detection using supervised machine learning. *International Journal of Advanced Computer Science and Applications*, 12(1). <https://doi.org/10.14569/ijacsa.2021.0120169>
- [15]Fuller, C. M., Biros, D. P., & Delen, D. (2011). An investigation of data and text mining methods for Real World Deception Detection. *Expert Systems with Applications*, 38(7), 8392–8398. <https://doi.org/10.1016/j.eswa.2011.01.032>
- [16]Garg, S., Gupta, S., & Gupta, B. (2022). Issues and challenges with fake reviews in digital marketing. 2022 International Conference on Computer Communication and Informatics (ICCCI). <https://doi.org/10.1109/iccci54379.2022.9740895>
- [17]Goswami, K., Park, Y., & Song, C. (2017). Impact of reviewer social interaction on online consumer review fraud detection. *Journal of Big Data*, 4(1). doi:10.1186/s40537-017-0075-6
- [18]Guha Majumder, M., Dutta Gupta, S., & Paul, J. (2022). Perceived usefulness of online customer reviews: A Review Mining Approach Using Machine Learning & Exploratory Data Analysis. *Journal of Business Research*, 150, 147–164. <https://doi.org/10.1016/j.jbusres.2022.06.012>
- [19]Hajek, P., Barushka, A., & Munk, M. (2020). Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Computing and Applications*, 32(23), 17259-17274. doi:10.1007/s00521-020-04757-2
- [20]Heydari, A., Tavakoli, M. ali, Salim, N., & Heydari, Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications*, 42(7), 3634–3642. <https://doi.org/10.1016/j.eswa.2014.12.029>
- [21]Jindal, N., & Liu, B. (2008). Opinion spam and analysis. *Proceedings of the International Conference on Web Search and Web Data Mining - WSDM '08*. doi:10.1145/1341531.1341560
- [22]Kumar, K. L., Desai, J., & Majumdar, J. (2016). Opinion mining and sentiment analysis on Online Customer Review. 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). doi:10.1109/iccic.2016.7919584
- [23]Manzoor, S. I., Singla, J., & Nikita. (2019). Fake news detection using machine learning approaches: A systematic review. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). <https://doi.org/10.1109/icoei.2019.8862770>
- [24]Martinez-Torres, M. R., & Toral, S. L. (2019). A machine learning approach for the identification of the deceptive reviews in the hospitality sector using unique attributes and sentiment orientation. *Tourism Management*, 75, 393–403. <https://doi.org/10.1016/j.tourman.2019.06.003>
- [25]Moon, S., Park, Y., & Seog Kim, Y. (2014). The impact of text product reviews on sales. *European Journal of Marketing*, 48(11/12), 2176–2197. <https://doi.org/10.1108/ejm-06-2013-0291>
- [26]Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2021). What yelp fake review filter might be doing? *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1), 409–418. <https://doi.org/10.1609/icwsm.v7i1.14389>
- [27]Nagi Alsubari, S., N. Deshmukh, S., Abdullah Alqarni, A., Alsharif, N., H. H. Aldhyani, T., Waselallah Alsaade, F., & I. Khalaf, O. (2022). Data Analytics for the identification of fake reviews using supervised learning. *Computers, Materials & Continua*, 70(2), 3189-3204. doi:10.32604/cmc.2022.019625

- [28]Narayan, R., Rout, J. K., & Jena, S. K. (2017). Review spam detection using opinion mining. *Advances in Intelligent Systems and Computing*, 273-279. doi:10.1007/978-981-10-3376-6_30
- [29]Noekhah, S., Fouladfar, E., Salim, N., Ghorashi, S. H., & Hozhabri, A. A. (2014). A Novel Approach for Opinion Spam Detection in E-Commerce. *Proc. of the 8th IEEE Int. Conf. on E-Commerce with Focus on E-Trus*, 1-8.
- [30]Rathor, A. S., Agarwal, A., & Dimri, P. (2018). Comparative study of machine learning approaches for amazon reviews. *Procedia Computer Science*, 132, 1552-1561. doi:10.1016/j.procs.2018.05.119
- [31]Saito, Y., & Klyuev, V. (2019). Classifying user reviews at sentence and review levels utilizing naïve Bayes. *2019 21st International Conference on Advanced Communication Technology (ICACT)*. doi:10.23919/icact.2019.8702039
- [32]Salminen, J., Kandpal, C., Kamel, A. M., Jung, S., & Jansen, B. J. (2022). Creating and detecting fake reviews of online products. *Journal of Retailing and Consumer Services*, 64, 102771. <https://doi.org/10.1016/j.jretconser.2021.102771>
- [33]Savage, D., Zhang, X., Yu, X., Chou, P., & Wang, Q. (2015). Detection of opinion spam based on Anomalous Rating Deviation. *Expert Systems with Applications*, 42(22), 8650-8657. doi:10.1016/j.eswa.2015.07.019
- [34]Shirsat, V. S., Jagdale, R. S., & Deshmukh, S. N. (2018). Sentence level sentiment identification and calculation from news articles using Machine Learning Techniques. *Advances in Intelligent Systems and Computing*, 371-376. doi:10.1007/978-981-13-1513-8_39
- [35]Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of sentiment reviews using N-gram machine learning approach. *Expert Systems with Applications*, 57, 117-126. doi:10.1016/j.eswa.2016.03.028
- [36]Tufail, H., Ashraf, M. U., Alsubhi, K., & Aljahdali, H. M. (2022). The effect of fake reviews on e-commerce during and after covid-19 pandemic: SKL-based fake reviews detection. *IEEE Access*, 10, 25555–25564. <https://doi.org/10.1109/access.2022.3152806>
- [37]Vimaladevi, K., & Dhanabhakaym, D. M. (2012). A study on the effects of online consumer reviews on purchasing decision. *Prestige International Journal of Management & IT - Sanchayan*, 01(01), 91–99. <https://doi.org/10.37922/pijmit.2012.v01i01.005>
- [38]Watson, F., & Wu, Y. (2021). The impact of online reviews on the information flows and outcomes of Marketing Systems. *Journal of Macromarketing*, 42(1), 146–164. <https://doi.org/10.1177/02761467211042552>
- [39]Zhang, D., Zhou, L., Kehoe, J. L., & Kilic, I. Y. (2016). What online reviewer behaviors matter? effects of verbal and nonverbal behaviors on detection of fake online reviews. *Journal of Management Information Systems*, 33(2), 456-481. doi:10.1080/07421222.2016.1205907

ANALYZE2GETHER.COM