SecuroGuard : E-Commerce Fake Review Detection Application

Project Proposal

Supervisor
Mr. Osama Ahmed Khan

Co-Supervisor
Mr. Maaz Ahmed

Submitted by

Faarah Khan
2578-2021

Hafsa Nisar
1988-2021

Department of Computer Science,
Hamdard University, Karachi.

2nd of July 2024

Table of Contents

Department of Computing

List of Tables

List of Figures

Department of Computing

1. Introduction

In today's online marketplace, the distinction between fake and counterfeit product reviews has blurred, creating a significant challenge for consumers seeking to make informed purchases. To solve this problem, application offers a powerful solution using advanced technology designed to detect and remove fake reviews. Application leverages advanced algorithms and machine learning capabilities to quickly scan more content, analyzing patterns and semantic clues that indicate fraudulent behavior. Character doing this not only protects the integrity of online feedback, but also increases trust in the authenticity of product reviews. This approach not only increases seller confidence by reducing misrepresentation, but also provides accurate information for the customer to make an informed decision and flexible designed to integrate seamlessly into existing e-commerce systems without impacting operational performance. This integration ensures that the use of application is both efficient and minimally disruptive, providing a great experience for both customers and merchants. By providing clear and reliable information, application creates a level playing field that allows businesses to thrive based on real customer feedback rather than reviews. Promotes a culture of transparency and accountability in the digital economy. By protecting the integrity of online reviews, application fosters trust between consumers and businesses, creating a foundation for growth and customer satisfaction. This trust is essential for long-term customer support and reputation, helping businesses succeed in a competitive online environment. By using technology to combat fraud and fraud, it not only protects customers from misinformation but also gives them the confidence to buy. Through performance measurement and a commitment to authenticity, application transforms online shopping into fair and transparent ecosystems where trust and confidence reign supreme. Finally, application promotes integrity and accountability helping to create a healthy digital economy where consumers and businesses safely share clear, reliable information.

2. Objective

To develop a system that detects and flags fake reviews using advanced machine learning techniques, thereby boosting online review credibility, increasing customer trust, and promoting fair competition among sellers.

3. Problem Description

In the rapidly growing e-trade panorama, the integrity of product opinions is critical for customer decisions. Alas, the upward push of fake evaluations compromises consideration and results in erroneous purchases.

- **What's the problem?** The core problem is identifying and eliminating these fake opinions to make certain clients get entry to reliable and honest data.

Department of Computing

- **Why is this important?** Fake evaluations lie to consumers and skew market opposition, unfairly reaping rewards positive merchandise and sellers. This no longer simplest negative aspects of genuine corporations that depend on sincere remarks however also erodes customer self-belief in e-trade structures, doubtlessly lowering usual online income and damaging platform credibility.

- **How will the problem be solved?** Our assignment proposes developing **SecuroGuard**, a gadget that employs system studying to analyze reviews for symptoms of inauthenticity. Through schooling on datasets of faux and genuine critiques, application will locate subtle variations and flag suspicious critiques. The gadget will provide actual-time analysis and remove questionable content, making sure customers get hold of only authentic comments. Moreover, application will offer special reports to customers and stakeholders, highlighting detected problems and movements taken. This could decorate belief in e-trade, assist truthful opposition, and maintain the integrity of assessment structures.

- **Benefits & Features:** Application will leverage superior algorithms to analyze overview styles and pick out fake content, offering only proper reviews to customers. The transition to automated detection necessitates cautious consideration of set of rules accuracy, records privateness, and system robustness. The device might be resilient to evolving fraud methods and capable of real-time analysis and reporting, making sure the credibility of online opinions.

4. Methodology

4.1. Project Approach

The "SecuroGuard" (E-commerce fake review detection application) aims to tackle fake reviews on e-commerce platforms using a spiral model approach. We start by defining clear goals and identifying potential risks. Then, we create prototypes and gather user feedback to refine our design. Next, we develop and test analytical models to distinguish genuine reviews from fake ones. Once the components are integrated, we rigorously test the system to ensure it works seamlessly. We also focus on clear communication to build user trust and continuously gather feedback for improvements. Regular updates and maintenance keep the system effective against evolving fraud tactics, ensuring it remains reliable and user-friendly.

Department of Computing

4.2. Why are we not using Agile or Scrum?

We choose not to use agile because our projects need to be precise and stable. This allows us to follow the waterfall model approach, where each phase begins with a plan with clear goals and achievement data.

4.3. Architecture & Design

4.3.1 Block Diagram
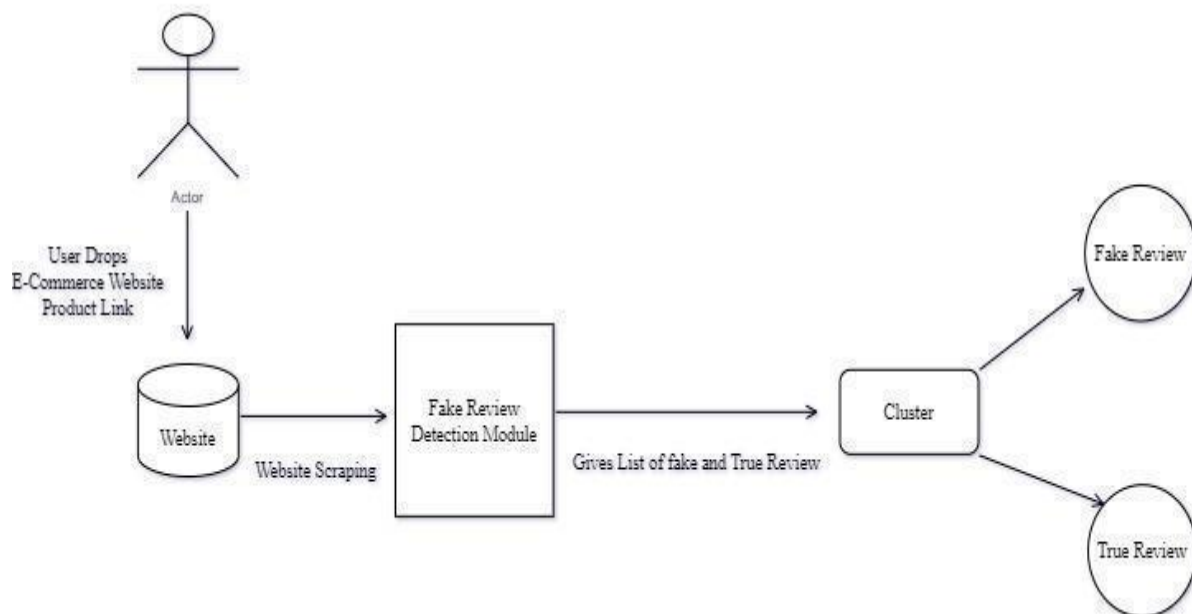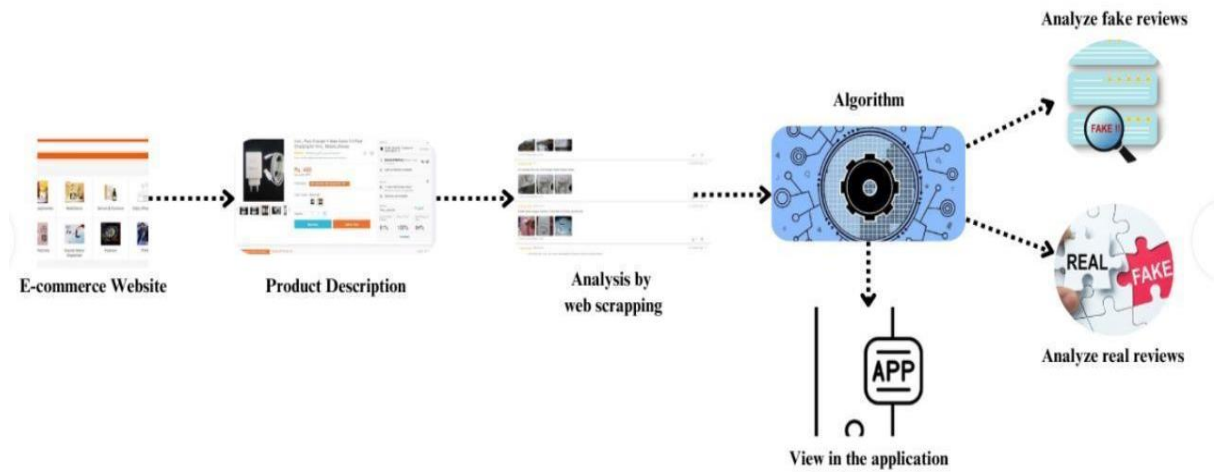

Figure 1: Block Diagram - Source (Visual Paradigm)

Department of Computing

4.3.2. Big Picture

Figure 2: Big Picture - Source (Canva)
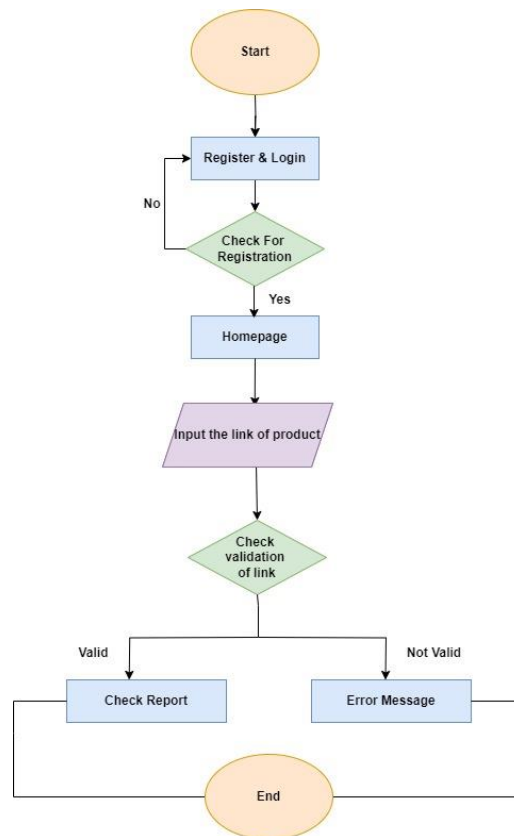
4.3.3. Flowchart

Figure 3: Flowchart - Source (Draw.io)

Department of Computing

5. Project Scope

The E-commerce Fake Review Detection Application effectively addresses the issue of fake reviews on e-commerce platforms by employing a rigorous approach to data collection and analysis. It begins by meticulously gathering information from diverse sources to ensure the accuracy and reliability of its findings. Utilizing advanced machine learning techniques and sophisticated natural language processing algorithms, the project identifies suspicious patterns indicative of fake reviews. This continual evolution of detection methods enables application to adapt and effectively examine reviews across various product categories, enhancing its ability to maintain integrity in online feedback.

Central to application solution is its user-friendly interface, designed to provide clear indicators of review authenticity and facilitate intuitive navigation for consumers. By presenting accurate information, the interface empowers buyers to make well-informed purchasing decisions aligned with their preferences and needs. Post-launch, the project prioritizes ongoing updates and regular maintenance to uphold stringent data protection standards and safeguard user privacy. These measures aim to foster a trustworthy online marketplace where sellers compete fairly and consumers rely on authentic reviews to guide their shopping experiences.

6. Feasibility Study

6.1 Risks Involved:

The project has some risks:

- First, the accuracy of our search algorithms may not be perfect, which can lead to inaccurate information. To solve this problem, we will use different data and continue to improve our algorithm.

- Second, people creating fake reviews may find new ways to avoid detection. To solve this problem, we will use a different model to learn and update new products. There are also data privacy risks and security issues, which we will address by following data protection laws and using strong encryption.

- Integrating application with different platforms can create challenges, so we need to make our systems flexible and compatible. Ultimately, limited access to computing power or human resources holds us back. We plan to solve this problem by carefully planning resources and using all tools. Access to real data from big data and e-commerce platforms is very important for learning. Finally, we will use data analytics platforms and machine learning technologies such as TensorFlow to effectively design and integrate application.

Department of Computing

6.2. Resource Requirement:

- We need high performance servers to meet demand. Instant data access from big data and e-commerce platforms is important for education. We also need skilled people like data scientists and software engineers

- Finally, we will use data analysis platforms as well as machine learning techniques such as TensorFlow to effectively design and integrate application.

7. Solution Application Areas

7.1 Is your project of some real value?

Yes, our Ecommerce Fake Review Detection application is of real value as it enhances the reliability of online product reviews, helping customers make informed purchasing decisions. By identifying and filtering out fake reviews, the application fosters a more trustworthy shopping environment, increasing customer confidence and satisfaction. This, in turn, benefits honest sellers by ensuring their products are evaluated fairly, leading to a more transparent and credible ecommerce marketplace.

7.2 Which industry or application domain are you targeting?

We are targeting the ecommerce industry, specifically online retail platforms where customer reviews play a crucial role in influencing purchasing decisions. This includes major online marketplaces, specially ecommerce stores, and consumer review websites. By integrating our Fake Review Detection application, these platforms can improve the authenticity of product reviews, enhancing the overall customer experience and fostering trust between buyers and sellers.

7.3 How may that target domain benefit from your solution?

- **Enhanced Trust:** By ensuring that reviews are genuine, customers can trust the feedback they read, leading to more confident purchasing decisions.
- **Improved Customer Experience:** A more reliable review system means customers are less likely to be misled by fake reviews, resulting in greater satisfaction with their purchases.
- **Fair Competition:** Honest sellers are protected from unfair competition by ensuring that fake positive or negative reviews do not distort their product ratings.
- **Reputation Management:** Ecommerce platforms can maintain a positive reputation by demonstrating a commitment to transparency and authenticity in customer reviews.
- **Increased Sales:** Trustworthy reviews can lead to higher customer engagement and increased sales, as customers feel more confident in their buying choices.

Department of Computing

8. Tools/Technology

The software tools utilized in the development of the "E-commerce Fake Review Detection" are listed below:

8.1. Tools

- Visual Studio code
- Anaconda
- Xampp
- Canva
- Visual Paradigm
- Draw.io
- MS Office Suite

8.2. Libraries

- Numpy
- Scikit-learn
- TensorFlow
- Pandas
- Matplotlib
- Flet

8.3. Programming Languages

- Python
- SQL

8.4. Version Control

- Git
- GitHub

Department of Computing

9. Responsibilities of the Team Members (RACI)

Project Deliverable Activity	Supervisor	Co Supervisor	Faarah Khan	Hafsa Nisar
Gather Requirements and Conducted Research	C, I	C, I	A	R
Requirements Locked	C, I	C, I	R	A
Collect Data	C, I	C, I	R	R
Learning and Skill Development	C, I	C, I	R	A
Web Scrapping Model	C,I	C,I	A	R
Model Development	C, I	C, I	R	A
Model Testing	C, I	C, I	A	R
FYP Report I	C, I	C, I	R	A
Develop Application	C, I	C, I	R	R
Integrate Model with Software	C, I	C, I	R	R
Integration Testing	C, I	C, I	R	R
FYP Report II	C, I	C, I	A	R

Table 2: RACI Chart – Source (MS Excel)

Department of Computing

10. Planning

10.1. Gantt Chart
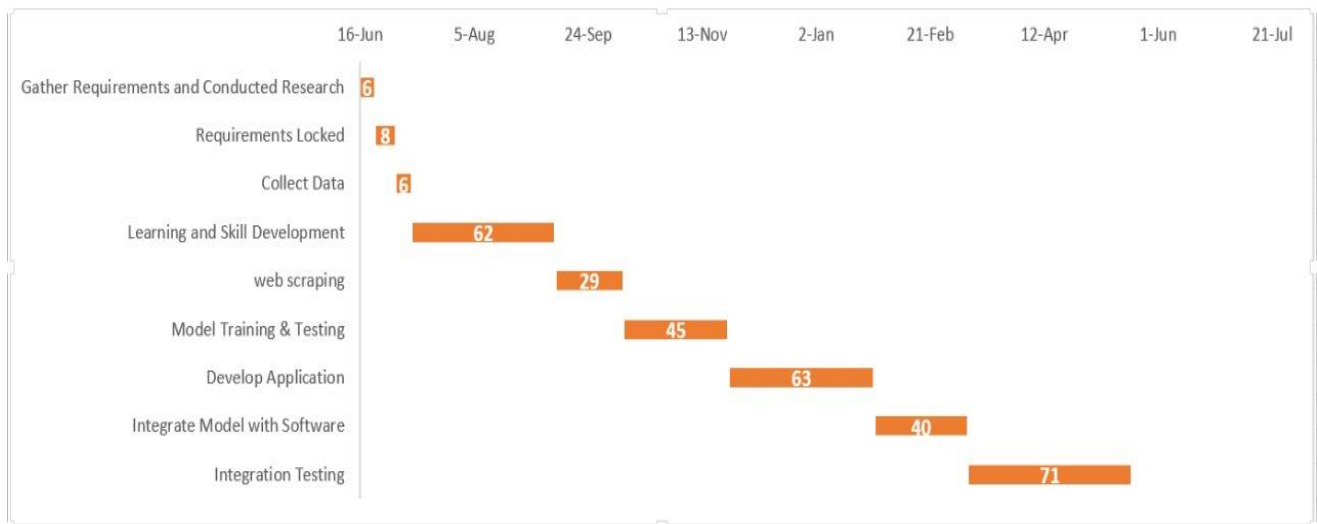
Figure 4: Gantt Chart - Source (MS Excel)

10.2. Milestones

10.2.1 First Evaluation

In the First evaluation following things will be done:

- ➢ Project Planning and Requirements Gathering
- ➢ Data Collection
- ➢ Learning and Skill Development
- ➢ Web Scrapping
- ➢ Model Development
- ➢ Model Testing

10.2.2 Second Evaluation

During the second evaluation phase, all of the anticipated functions of the entire module, such as

- ➢ Mobile Application Development
- ➢ System Integration
- ➢ Testing

Department of Computing

10.3. Budget/Costing

S No	Product Name	Price
1	HP 15 series laptop	120,000
2	Utility Expense	32,000
3	Development Cost	253,000
4	Testing Cost	35,000
Total		**440,000**

Table 3: Budget/Costing - Source (MS Excel)

11. References

[1] Fake review detection in e-Commerce platforms using aspect-based sentiment analysis. Journal of Business Research. November 2023, Volume 167, Pages 114143. Petr Hajek, Lubica Hikkerova, Jean-Michel Sahut. Accessed June 30, 2024. Available at: https://doi.org/10.1016/j.jbusres.2023.114143.

[2] Fake review detection system for online E-commerce platforms: A supervised general mixed probability approach. Decision Support Systems. December 2023, Volume 175, Page 114045. Jiwei Luo, Jian Luo, Guofang Nan, Dahui Li. Accessed June 30, 2024. Available at: https://doi.org/10.1016/j.dss.2023.114045.

[3] DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network. Information Processing & Management. March 2018, Volume 54, Issue 2, Pages 255-268. Wen Zhang, Yuhang Du, Taketoshi Yoshida, Qing Wang. Accessed June 30, 2024. Available at: https://doi.org/10.1016/j.ipm.2018.03.007.

[4] Kumar, A., Gopal, R. D., Shankar, R., Tan, K. H. (2022). Fraudulent review detection model focusing on emotional expressions and explicit aspects: investigating the potential of feature engineering. Decision Support Systems, Volume 155, April 2022, 113728. Accessed June 30, 2024. Available at: https://doi.org/10.1016/j.dss.2021.113728.

[5] Opinion spam detection by incorporating multimodal embedded representation into a probabilistic review graph. November 13, 2019. Neurocomputing. Last accessed June 30, 2024. Authors: Liu Yuanchao, Pang Bo, Wang Xiaolong. https://doi.org/10.1016/j.neucom.2019.08.013

[6] A deep learning approach for detecting fake reviewers: Exploiting reviewing behavior and textual information. March 2023. Decision Support Systems. Last accessed June 30, 2024. Dong Zhang, Wenwen Li, Baozhuang Niu, Chong Wu. https://doi.org/10.1016/j.dss.2022.113911

[7] Fake online reviews: Literature review, synthesis, and directions for future research. May 2020. Decision Support Systems. Last accessed June 30, 2024. Yuanyuan Wu, Eric W.T. Ngai, Pengkun Wu, Chong Wu. https://doi.org/10.1016/j.dss.2020.113280

[8] Fraud detection in online consumer reviews. 2010. Decision Support Systems. Last accessed July 1, 2024. Nan Hu, Ling Liu, Vallabh Sambamurthy. https://doi.org/10.1016/j.dss.2010.08.012

[9] Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus. September 2016. Journal of Retailing and Consumer Services. Last accessed July 1, 2024. Andreas Munzel. https://doi.org/10.1016/j.jretconser.2016.06.002

[10] Detection of review spam: A survey. 1 May 2015. Elsevier. 1 July 2024. Atefeh Heydari, Mohammad Ali Tavakoli, Naomie Salim, Zahra Heydari. https://doi.org/10.1016/j.eswa.2014.12.029.

11.1 Literature Review

Literature Review - Summary

#	Paper Title	Publication Year	Publisher	Source	Keywords	Gap	Reference
1	Fake review detection in e-Commerce platforms using aspect-based sentiment analysis	2023	Journal of Business Research	Google Scholar	Fake review detection, Aspect-based sentiment analysis, E-commerce platforms, Machine learning, Natural language processing	This study addresses the differences between existing fake audits by focusing on dimension-based reasoning (ABSA), which identifies differences between products/services reviewed online to improve truth discovery. This approach makes a difference where previous methods often treated reviews as if they felt the same across all dimensions, ignoring nuances that could indicate fake reviews.	[1]
2	Fake review detection system for online E-commerce platforms: A supervised general mixed probability approach	2023	Decision Support Systems	Science Direct	Artificial intelligence in "smart cars"	This research focuses on the development of a new truth management method (FRDGMPM) to detect fake reviews using the differences between the characteristics of non-fake products and fake reviews. It expands existing communication, behavior and relationship processes into integrated processes with the aim of improving reality, especially when Initial examination information is limited.	[2]
3	DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network	July 2018	Information Processing & Management	Science Direct	Deceptive review identification, recurrent convolutional neural network (RCNN), deep learning, word context, online reviews, fraud detection	This article addresses the differences between existing automated fraud detection methods by proposing the DRI-RCNN method, which uses single-word concepts and deep learning (especially RCNN) to improve the accuracy of distinguishing fake from real analysis. This method works better than traditional methods by taking into account contextual information in the analysis, which is important to determine the truth of online fraud platforms.	[3]
4	Fraudulent review detection model focusing on emotional expressions and explicit aspects: investigating	April 2022	Decision Support Systems	Science Direct	Fraudulent review detection, feature engineering, machine learning, NLP, text mining, M-SMOTE,	The goal of this article is to increase fraud detection by emphasizing the use of technology to eliminate "review-centric" and non-"reviewer-centric" ones. It solves problems such as skewness and distribution inequality in data distribution and proposes a new M-SMOTE model to improve detection accuracy. This approach makes a difference by combining best practices with pre-existing information strategies to improve fraud detection as spammers change their strategies on online platforms.	[4]

Department of Computing

	the potential of feature engineering				data pre-processing, skewed distributions, e-commerce	
5	Opinion spam detection by incorporating multimodal embedded representation into a probabilistic review graph	November,, 2019	Neurocomputing	Science Direct	Opinion spam detection, multimodal embedded representation, probabilistic review graph, neural network with attention mechanism, heterogeneous graph model, Markov Random Fields (MRFs), textual features, rich features, fake review detection, Dianping dataset, China	This paper presents a new method combining multimodal representation learning and modeling techniques to solve the challenge of spam detection in online medicine. Improving spam detection accuracy in big data highlights the importance of determining the relationship between different entities (reviews, articles, products) to fill gaps in the use of contextual elements and rich metadata features. This study demonstrates the significant development of culture-based classification using the image model as a possibility of changing belief through the media in the coordination of advocacy organizations.	[5]
6	A deep learning approach for detecting fake reviewers: Exploiting reviewing behavior and textual information	2023	Decision Support Systems	Science Direct	Fake reviewer detection, deep learning, reviewing behavior, textual information	This paper presents a deep learning-based framework that uses behavioral recommendations and texts to solve the challenge of detecting fake reviewers on online platforms. It aims to reduce the number of people working on the infrastructure and increase the detection accuracy compared to traditional methods.	[6]
7	Fake online reviews: Literature review, synthesis, and directions for future research	2020	Elsevier	Science Direct	Fake reviews, e-commerce, literature review, antecedent-consequence-intervention framework, future research directions	This document states that extensive research was not done to review and collect previous content, resulting in fake online reviews. It presents previous intervention methods and results and identifies future research questions and recommendations that will guide further research in this area. Additionally, it highlights the paucity of excellent data available for examining fake reviews, which remains a challenge in this field.	[7]
8	Fraud detection in online consumer reviews	2011	Elsevier	Google Scholar	Online consumer reviews, review fraud,	This article examines the most common forms of review fraud in online consumer reviews and specifically examines how sellers, publishers, and authors use these review techniques to influence	[8]

Department of Computing

				manipulation, product quality, Amazon, Barnes & Noble	customers' opinions. Explores the relationship between review management and product quality, showing that management reduces the credibility and information of online reviews. This study also explores consumers' awareness and its ability to reduce management's influence on purchasing decisions. Additionally, it highlights differences in fraud by comparing management levels across different platforms (Amazon and Barnes & Noble).	
9	Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus	2016	Elsevier	Science Direct	Fake reviews, identity disclosure, consensus, online reviews, consumer trust, deception detection	This study investigates how information about reviewer identity and approval of reviews affects consumers' ability to detect fake reviews. He talked about the increasing number of fake reviews online and the impact this has on consumer trust in review platforms. This study supports this by evaluating the effectiveness of cues (self-report and confirmation) to help consumers distinguish genuine reviews from fake reviews. It also highlights the importance of consumer education and awareness in reducing the negative impact of fake reviews on online reviews.	[9]
10	Detection of review spam: A survey	2015	Frontiers in Psychology	Google Scholar	Review spam, spam detection techniques, individual spammers, group spammers, opinion mining, sentiment analysis	This study examines and classifies various methods for identifying spam comments; distinguishes between methods that target spam, individual spammers, and group spammers. It analyzes the advantages and disadvantages of these technologies, providing insight into their accuracy, effectiveness, and usability in various contexts. This study aims to provide guidance to practitioners and researchers in the field of fraud detection and to address the growing problem of fraud associated with online review platforms.	[10]

Department of Computing