Hamdard University

Department of Computing

Final Year Project

SecuroGuard

E-Commerce Fake Review Detection Application

FYP-011/FL24

Software Requirements Specifications

Submitted by

Faarah Khan (2578-2021)
Hafsa Nisar (1988-2021)

Supervisor

Mr. Afzal Hussain

Co-Supervisor

Mr. Maaz Ahmed

2025

Faculty of Engineering Sciences and Technology

Hamdard Institute of Engineering and Technology

Document Information

Project Title	SecuroGuard E-Commerce Fake Review Detection Web-App
Project Code	FYP-011/FL24
Document Name	Software Requirements Specifications
Document Version	2.0
Document Identifier	FYP-011/FL24-SRS
Document Status	~~Draft~~/ Final
Author(s)	Faarah Khan Hafsa Nisar
Approver(s)	Mr. Afzal Hussain Mr. Maaz Ahmed
Issue Date	17-01-2025

SecuroGuard E-Commerce Fake Review Detection Web-App	Version:2.0
Software Requirements Specifications	Date: 17/01/2025
FYP-011/FL24	

Name	Role	Signature	Date
Faarah Khan	Team Leader		17-01-2025
Hafsa Nisar	Team Member 2		17-01-2025
Mr. Afzal Hussain	Supervisor		17-01-2025
Mr. Maaz Ahmed	Co-Supervisor		17-01-2025
Mr. Faheem Ahmed Khan	Project Coordinator		17-01-2025

Revision History

Date	Version	Description	Author
12/12/2024	1.0	First Draft	Faarah Khan, Hafsa Nisar
17/01/2025	2.0	Second Draft	Faarah Khan, Hafsa Nisar

Definition of Terms, Acronyms, and Abbreviations

Terms	Description
SRS	Software Requirement Specification: A document that defines the functional and non-functional requirements of the software system.
API	Application Programming Interface: A set of protocols and tools that allows different software applications to communicate with each other.
HTTP	Hypertext Transfer Protocol: A protocol used for transferring web pages on the internet.
HTTPS	Hypertext Transfer Protocol Secure: An extension of HTTP that ensures secure communication by encrypting data with SSL/TLS.
ML	Machine Learning: A branch of AI that allows computers to learn and make decisions based on data without explicit programming.
NLP	Natural Language Processing: A field of AI that focuses on enabling computers to understand and process human languages.
UI/UX	User Interface / User Experience: UI refers to the design and layout of user interfaces, while UX is focused on the overall experience of the user interacting with the system.
SSL	Secure Socket Layer: A security protocol that provides encrypted communication between web servers and browsers.
URL	Uniform Resource Locator: The address used to access resources on the web.
AI	Artificial Intelligence: The simulation of human intelligence in machines designed to think and act like humans.
Fake Reviews	Reviews that are intentionally misleading or deceptive, often posted to promote a product or service or discredit competitors.
Sentiment Analysis	A technique in NLP that analyzes and determines the sentiment (positive, negative, or neutral) expressed in text data, such as reviews.
Text Mining	The process of extracting useful information from unstructured text data, often used in fake review detection to find patterns or anomalies in reviews.

Terms	Description
Anomaly Detection	A technique used to identify patterns in data that do not conform to expected behavior, often used to identify fake reviews based on unusual review patterns.
Review Spam	Reviews that are irrelevant or repetitive, often posted in large numbers to manipulate the product or service's rating or reputation.
Reviewer Profile	A profile of a user who posts reviews, including their posting history and behavior, used to detect suspicious or fake review patterns.
Crowdsourcing	The practice of obtaining data or input by soliciting contributions from a large group of people, often used in verifying the authenticity of reviews.
Authorship Attribution	The process of determining the author of a piece of text, which can help identify whether reviews come from the same or suspicious authors.
Review Verification	Techniques used to verify the authenticity of a review, such as checking for inconsistencies or using machine learning models to assess likelihood of fakeness.
Machine Learning Model	A mathematical model used to analyze and predict data trends; in fake review detection, it's used to predict whether a review is genuine or fake based on patterns in data.
Deep Learning	A subset of machine learning involving neural networks with many layers, useful in detecting complex patterns in data such as fake review behavior.
Feature Extraction	The process of identifying and selecting relevant features (like sentiment, keywords, or reviewer behavior) for use in training machine learning models for fake review detection.
Classification Algorithm	An algorithm that categorizes data into predefined groups, such as identifying reviews as real or fake, based on features extracted from the text.
Data Labeling	The process of tagging data (e.g., reviews) with predefined labels, such as "fake" or "real," which is used to train machine learning models.
Social Proof	The concept that people are influenced by others' opinions and reviews, which can be manipulated in fake review detection systems.
Review Velocity	The rate at which reviews are posted for a product or service, used to detect suspicious activity or sudden bursts of fake reviews.
Reviewer Behavior Analysis	Analyzing a reviewer's activity and history to detect patterns that might indicate fake or paid reviews, such as reviewing multiple products in a short time.

Table of Contents

1. Introduction

Fake reviews have become a major problem in the digital economy, undermining customer confidence and interfering with fair competition. This problem is addressed by the E-commerce Fake Review Detection Application, which uses sophisticated machine learning algorithms to recognize and flag false reviews. This program enhances the legitimacy of online reviews and fosters confidence between customers and companies by integrating seamlessly with current ecommerce systems. It provides companies with trustworthy, transparent information while assisting consumers in making educated purchase decisions. This strategy promotes an equitable and responsible digital economy, protecting the legitimacy of internet reviews and guaranteeing a level playing field for all parties.

1.1 Purpose of Document

The purpose of this paper is to provide a clear description of the functional and non-functional requirements for the application that defects fake reviews in e-commerce. Through real-time detection and reporting of fraudulent reviews, the program is intended to confirm the legitimacy of reviews on e-commerce platforms. It provides a thorough knowledge of the system's goal, scope, and operational context for the development team, academic evaluators, and future stakeholders. The application aims to improve online purchasing transparency, promote fair competition, and foster customer trust by tackling the problem of fraudulent reviews.

1.2 Intended Audience

The project team, which consists of two people working together to build the E-Commerce Fake Review Detection Application, is the main audience for this document. Academic advisers and assessors who are in charge of analyzing the project's plan, development and results can also use it as a guide. The paper also outlines potential interactions with stakeholders, including consumers, ethical merchants, and administrators of e-commerce platforms, all of whom stand to gain from a more open and reliable review process. This material lays the groundwork for future scholarly inquiry or perhaps industrial implementation.

2. Overall System Description

2.1 Project Background

Consumer selections are greatly influenced by customer reviews, which have become an essential component of internet buying. However, the increasing number of fraudulent reviews has damaged consumer trust and interfered with the fairness of the market. The goal of this project is to develop a potent tool that efficiently detects and gets rid of fraudulent reviews using cutting-edge machine learning and natural language processing approaches.

2.2 Problem Statement

On E-commerce sites, fake reviews are a serious problem as they mislead consumers and stop honest competition. Product evaluations are essential for assisting customers in making wellinformed selections, but the rise in fraudulent reviews erodes consumer confidence and frequently results in poor-quality purchases. These false reviews damage ecommerce platforms' reputations, undermine consumer trust, and offer particular goods and vendors unfair advantages. To solve this, a trustworthy method for spotting and eliminating fraudulent reviews is required, guaranteeing that consumers can rely on the opinions they read and that online purchasing stays trustworthy and equitable.

2.3 Project Scope

The E-commerce Fake Review Detection Application allows users to paste the link of any product into a webpage and check for reviewers accuracy, thus eliminating the growing issue of fake reviews. Electricity bills can be paid efficiently by utilizing fraud detection patterns in machine learning and natural language processing. That said, tablets, desktops, and mobile devices can all access the application with the same ease because of the responsive design. The application guarantees that flagged reviews can be easily understood by its users by providing ample visual representations and information so that their purposes could be clearer. The application provides consistent methods in tackling fake reviews, so it effectively creates a secure environment that is trustworthy.

2.4 Not In Scope

This project does not include the monitoring of reviews that the project members have flagged or responding to fake reviews on non-e-commerce platforms like blogs or social media. Although the potential scope is to detect fake reviews, the framework can never be fully accurate as scammers will always come up with new ways to manipulate the said framework. This encompasses the authentication of the reviewers and prevention of scammers, but not conquering fraudsters through legal means. Furthermore, since the primary focus of the project is to build a web application, Internet application design will be omitted.

2.5 Project Objectives

The purpose of this project is to develop a web application that employs advanced natural language processing and machine learning techniques to detect and flag fake reviews. The aim of the strategy is to enhance the authenticity of the online reviews by ensuring that all the customers are presented with real and more reliant reviews. This will help to reduce the effect of fake reviews, hence making it easier for contestants to compete and consumers to get trust. The ultimate ambition of this project is to create an honest and honest platform that endorses a healthy and trustworthy online economy.

2.6 Stakeholders & Affected Groups

The students entrusted with this particular project along with the faculty advisors helping in the assessment are the primary stakeholders. Also, the e-commerce administrators who will aid in implementing the solution, and the discerning customers trusting constructive reviews will become important stakeholders at a later stage. This model will be beneficial also to fair vendors who would otherwise be harmed by the existence of false reviews, as it ensures healthy competition and helps protect their image.

2.7 Operating Environment

As the application will be a web-based system, it will be compatible with popular web browsers such as Chrome, Firefox, Safari, and Edge. It will also be built to integrate easily via APIs with existing systems such as e-commerce platforms to enable seamless data transmission between systems. The application will be compatible with the desktop, laptop, tablet, and smartphone which means users will be able to access and use the system from practically anywhere and any device.

2.8 System Constraints

The effectiveness of the system would also require the availability of annotated datasets for the training of the machine learning model, and such a system would have to comply with data protection laws to ensure the privacy of its users. For instance, when developing the system, there are several factors to consider such as the available resources of the system, which are likely to affect the speed of the analysis of data. Additionally, there is a need to provide crossplatform support for different e-commerce platforms which may pose integration and customization challenges.

2.9 Assumptions & Dependencies

In order for this project to be successful, a number of assumptions and dependencies will have to be made. First of all, in order for the machine learning algorithm to be trained, the system will have to be provided with comprehensive and trustworthy information, which is important for effective fake review detection. Furthermore, reliable internet connectivity will be needed as well for data analysis and processing. One more notable requirement is the interest of ecommerce platforms in integrating the solution as that would determine the reach of the deployed system and the likely impact on the market.

3. External Interface Requirements

3.1 Hardware Interfaces

The program will have interaction within standard consumer electronics comprising desktops, laptops, tablets and smartphones as well as high powered web servers responsible for data processing, analysis and storage. These servers will facilitate seamless work on a range of devices and ensure efficient real-time detection.

3.2 Software Interfaces

For efficient and effective storage solutions, the application will integrate Firebase databases to facilitate storage as well as other site APIs for data reviews. To categorize the reviews, a BERT AI model will be used. The HTML, CSS, and JavaScript technologies will be utilized to develop the front-end interface to the application.

3.3 Communications Interfaces

The system is expected to implement WebSockets to provide the live update status on the review changes, email services to send out alerts, RESTful APIs to interact with as well as integrate to the e-commerce platforms, and HTTP/HTTPS to enable secure communication channels.

4. System Functions / Functional Requirements

4.1 System Functions

This section details the major and extra functionalities that are provided by the Fake Review Detection System. These functions have some constraints concerning programming language, interface, design and performance. They are divided into three categories: Evident, Hidden, and Frill. The functions are described in a way that ensures that traceability and clarity are preserved.

4.1.1 Function Categories

Ref #	Functions	Category	Attribute	Details & Boundary Constraints
R1.1	Permit users to register or sign in with Google	Evident Usability	Authentication	Must integrate with Google OAuth for authentication.
R1.2	Display the landing page with an introductory overview	Evident Responsiveness	Performance	Should load efficiently and provide a responsive user experience.
R1.3	Provide an Analyze Page for review analysis	Evident Interaction	User Interface	Must display a clear interface for entering product URLs.

R1.4	Scrape reviews from the provided product URL	Hidden Performance	Efficiency	Review scraping should be efficient and seamless.
R1.5	Detect fake reviews and generate analysis	Hidden Accuracy	AI Model	The model's accuracy in detecting fake reviews must be at least 90%.
R1.6	Display the review analysis report	Evident Experience	Insights	Should include percentage of fake reviews, graphs, and other insights.
R1.7	Flag fake reviews for further action	Hidden Security	Accountability	Logs all flagged reviews for transparency and tracking.
R1.8	Provide a Contact Us page	Evident Communication	Support	Must include a form for inquiries and an FAQ section.
R1.9	Offer How It Works guidelines	Evident Onboarding	User Guidance	Should include step-by-step instructions for using the system.

4.1.2 Verification of System Functions

- Users must also have an option to sign up or log in using their Google account using this system.
- The system should display a landing page with an introduction.
- The system should also have an Analyze Page for review analysis.
- The system should be able to scrape reviews from the provided product's URL.
- The system must be able to detect fake reviews and give a report of analysis.
- The system should allow for the generation of a review analysis report.
- The system should improve its functionality by allowing for the flagging of fake reviews for action.
- There should be a Contact Us page where users can go to ask their questions within the system.
- The system should include a How It Works feature that will help new users get started.

4.1.3 Restrictions on Design

- Web possible solutions, such as developing with HTML/CSS, JavaScript, and also interfacing with Firebase backend services can be used in the development of the system.
- When making the system, the developers must adhere to responsive design rules so that it can be used on PCs, tablets, and smartphones.
- Scalability is paramount: The systems need to control an increasing number of users and their interactions and review as it grows in size.

4.1.4 Interface Requirements

- **Sign In Page:** Google OAuth service to check the user.
- **Analyze page interface:** Input box for the product's URL, progress bar, analysis walkthrough.
- **The results page interface:** Main points and percentages of fake reviewers are exhibited graphically.
- **Contact Us Page:** contact address for help, frequently asked questions, and a question-asking form.

System Attributes / Nonfunctional Requirements

Attribute	Details and Boundary Constraints	Category
Response Time	All user-facing interactions, such as login and navigation, should be efficient.	Mandatory
Concurrent User Load	The system must support a minimum of 100 users connected simultaneously.	Mandatory
Ease of Use	The interface should be user-friendly and intuitive, requiring minimal training.	Optional
Security	All user data and transactions must be encrypted using industry-standard protocols.	Mandatory
System Uptime	The system must maintain 99.9% uptime annually.	Mandatory
Scalability	The system must handle up to 10,000 reviews daily without performance degradation.	Optional

4.2 Use Cases

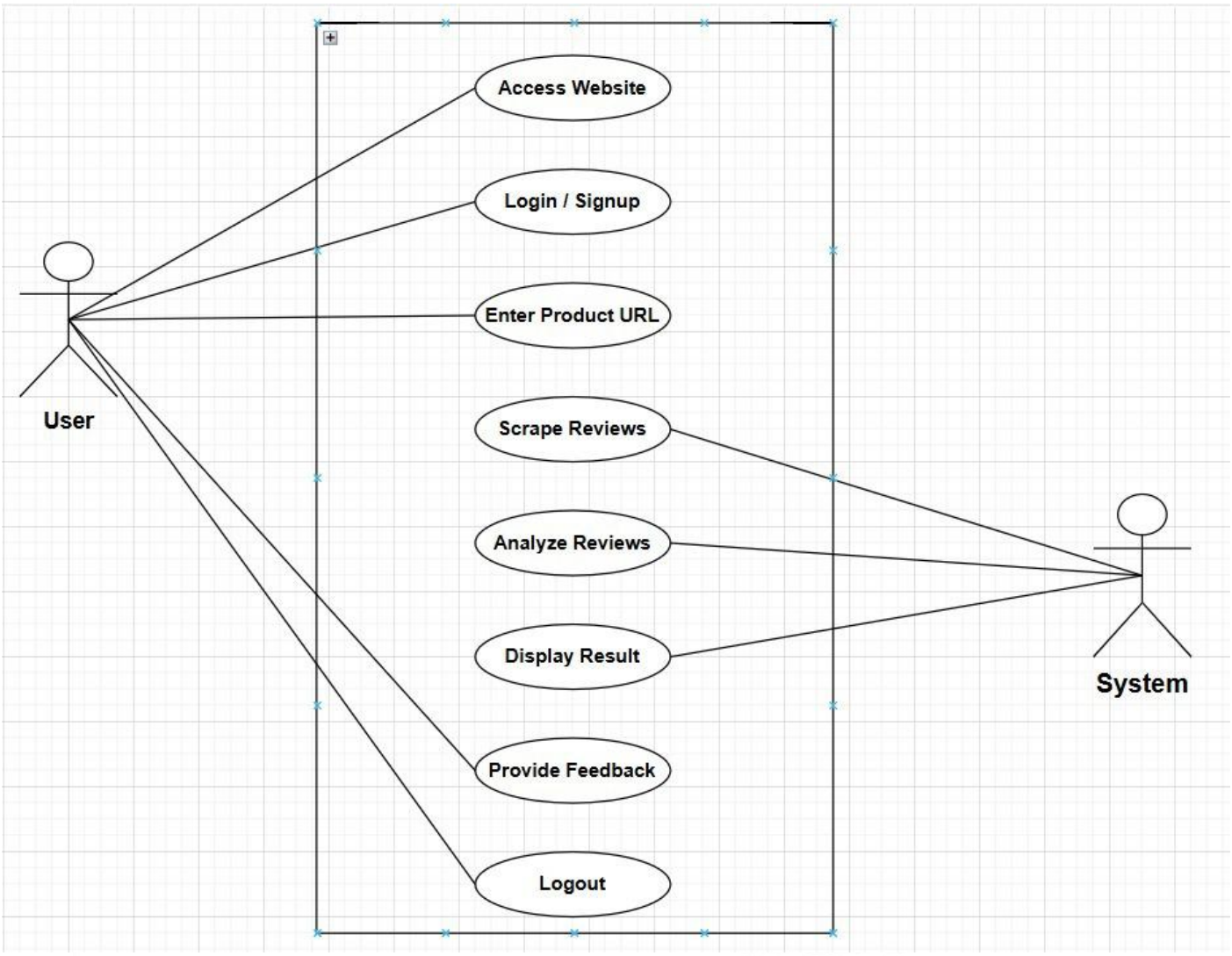
4.2.1 List of Actors

- **User:** The actor operates the system through the site, signs into the platform, submits the URLs for products, and then observes the output generated.
- **Service:** The service accumulates evaluations through its back end, performs the analysis, and then presents the data to the user in an understandable format.

4.2.2 Use Case List:

- **Go to the site:** To interface with the system, the user has to access the site.
- **Login/Signup:** A user of the system is required to log in or sign up with the system through Google authentication.
- **Provide the Product's Link:** The user will provide the product's link to carry out a review analysis.
- **Extract Reviews:** Based on the product's URL uploaded, the system will extract the reviews.
- **Review Analysis:** The reviews will be evaluated by the system in order to find fake ones.
- **Results Presentation:** The user will be able to view the results of the analysis as well as insights and charts generated by the system.
- **User's Feedback:** The user has to provide feedback with regard to the analysis conducted or the performance of the system.
- **Sign out:** The user has to sign out from the system in order to end the session.

4.2.3 Use Case Diagram:



4.2.4 Description of Use Cases:

Attribute Details	Details
Use Case Name	Detect Fake Reviews
Actors	User, System/Server
Purpose	To identify fake reviews based on submitted product URLs and analyze reviews for authenticity
Pre-Conditions	The user must be logged into the application and have submitted a valid product URL
Post-Conditions	The system displays the analysis results, highlighting flagged fake reviews
Normal Flow	1. Detect Fake Reviews 2. The user reasons to enter the product link 3. And then the 'Analyze' button 4. Evaluates and begins recording the reviews 5. Any result that is flagged is shown by the system 6. Finally, user exists the application

Typical Process: Detecting Fake Reviews

Actor Action	System Response
The use case initiates once the user enters their credentials in the application.	The application displays the home screen upon user verification by the system.

2. The user inputs the link of the product after clicking on the “Analyze Reviews” button.	The system verifies the link provided and prepares itself for review scraping.
3. The user selects the 'Analyze' button.	The system scrapes reviews from a specified URL.
4. To determine the legitimacy of reviews, the system analyzes the reviews provided by users.	Generated results include the flagged reviews that were identified as fake.
5. The system displays the results of the analysis.	The user receives a comprehensive report with visuals.
6. The user of the application signs out.	The session is terminated, and the system redirects to the login page.

5. Non - Functional Requirements

5.1 Performance Requirements

This is designed to ensure that reviews are processed and flagged instantly as they come. In this sense, it provides users with the ability to generate review analysis reports instantly as it eliminates any delays. On top of that, the system is highly stable and quite resilient as it can handle a minimum of 100 concurrent users without degradation in performance.

5.2 Safety Requirements

The system has been designed to ensure the integrity of all actions performed by the system and its users. Regardless of any unexpected scenarios, the system is developed to automatically save and store every single piece of executed data. When the system restarts after a crash, it restarts from the latest save, thereby ensuring that no data is lost. Moreover, no review that has been flagged can be deleted without system permission which further prevents the accidental deletion of critical information and ensures that it can be accurate and complete.

5.3 Security Requirements

For enhanced user security, all data transmitted through user accounts and stored on a system should be encrypted. Additionally, to achieve a higher level of security within system accounts, Multi-Factor Authentication (MFA) should be put in place. Unsuccessful login attempts should trigger a system lock when the figure of such attempts rises to three. Lastly, to prevent unassociated users from altering fraud detection settings, only authenticated accounts will hold the permissions that are required to update the detection model.

5.4 Reliability Requirements

In order for the system to be able to quickly recover from a potential mishap, it is required that 99.9 uptime is maintained while an automatic backup for the server is created once every hour. Furthermore, it has to be designed in a way that is able to handle up to three server outages simultaneously without loss of availability and performance of the hardware.

5.5 Usability Requirements

Laws of responsive design must be employed in the development of the interface so that it works seamlessly across the different platforms; that is desktop, tablet, and mobile device. The features should be intuitive enough not to take more than three clicks to access an aspect of the interface. New users will also be able to use a “How It Works” module that will simplify the procedure of signing up for them.

5.6 Supportability Requirements

The system should guarantee the deployment of zero-downtime which will permit the installation of patches and conduct upgrades without affecting the ongoing services. Due to its modular design, it must be easy to connect with other systems such as CRM or e-commerce. In addition, it must also be easy to upgrade the detection model to leverage any advancement made in relation to artificial intelligence.

5.7 User Documentation

An in-depth 'How It Works' section should be provided, which will be highlighting the main functionalities of the system, such as how to register or log in, how to search and read product reviews, and how to interpret the review analysis reports. Functions of the system such as accepting or rejecting reviews, and the updating of detection models should also be included in this section. There should also be a web based frequently asked questions (FAQ) in order to help users find answers to their questions and such documentation should include contact information for technical assistance.

6. 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.