 University of South Wales

Prifysgol De Cymru

Faculty of Computing, Engineering and Science

MSc Project

Smart Product Hunting for Facebook and Craigslist

Supervisor

Dr. Janusz Kulon, Associate Professor (Research and Development)

Faculty of Computing, Engineering and Science

Student Name: Ibtehaj Gulzar

Enrolment number: 30083149

Course: MSc Computer Science

Module ID: CS4T702

Module Title: MSc Project

September 2023/24

UNIVERSITY OF SOUTH WALES

PRIFYSGOL DE CYMRU

FACULTY OF COMPUTING, ENGINEERING & SCIENCE

SCHOOL OF COMPUTING & MATHEMATICS

# Statement of Originality

This is to certify that, except where specific reference is made, the work described in this project is the result of the investigation carried out by the student, and that neither this project nor any part of it has been presented, or is currently being submitted in candidature for any award other than in part for the MSc award, Faculty of Computing, Engineering and Science from the University of South Wales.

Signed Ibtehaj Gulzar………………………………(Student)

Date 19/09/2024………………………………………………….

**Abstract**

This study explores the emerging markets of e-commerce and, more specifically, smart shopping strategies on social media, specifically Facebook marketplace and Craigslist. With these platforms playing a more prominent role in the eCommerce sphere, so has the demand for better algorithms, data mining, and machine learning for the purpose of improving the search and recommendations of products. As a social networking site turned all-in-One e-commerce platform, Facebook uses user data to enhance the pitch, suggest alternatives, and fetter a buyer’s enjoyment in the process. On the other hand, traditional classified portal Craigslist has its limitations due to its simple design and low level of innovative search options. It exposes how the discovery of products can be improved by using ontology-based approaches, natural language processing, and sentiment analysis, and how issues of fraud and users’ security can be solved. Moreover, it does look into the application of artificial intelligence in the enhancement of user experience, promotion and marketing, and customers’ satisfaction. The application of AI tools, like the Random Forest algorithm, is considered in the context of identification of product trends and optimization of e-commerce planning. This research stresses the necessity of applying the AI big data technology to improve the smart product hunting functions that benefit digital platforms to enhance users’ efficient e-commerce experience.

**Acknowledgments**

I would like to express my profound gratitude to Allah Almighty for His blessings and guidance throughout the course of this research. I also extend my heartfelt thanks to my parents for their unwavering support and prayers. My sincere appreciation goes to my Supervisor, Teachers Departments for providing the academic environment and resources necessary for this research.

**List of Figures**

[Figure ‎1‑1. Project Timeline 8](#_Toc176172292)

[Figure ‎2‑1 Proposed summarization model (Hou et al., 2019) 10](#_Toc176172293)

[Figure ‎2‑2 Workflow of Research (Abighail et al., 2023) 12](#_Toc176172294)

[Figure ‎2‑3 Recommendation system model (Xu and Sang, 2022) 15](#_Toc176172295)

[Figure ‎2‑4 Random forest model (Xu and Sang, 2022) 15](#_Toc176172296)

[Figure ‎2‑5 Most Popular Social Networking Sites (Singh and Singh, 2018) 17](#_Toc176172297)

[Figure ‎2‑6 @Samsung Mobile Followers across (Singh and Singh, 2018) 17](#_Toc176172298)

[Figure ‎2‑7 The general process of sentiment analysis (Xu, Chang and Jayne, 2022) 18](#_Toc176172299)

[Figure ‎2‑8 Methodology used (S et al., 2023) 19](#_Toc176172300)

[Figure ‎3‑1.Routes 29](#_Toc176172301)

[Figure ‎4‑1. Backend Structure 34](#_Toc176172302)

[Figure ‎4‑2. USER REGISTRATION AND AUTHENTICATION 35](#_Toc176172303)

[Figure ‎4‑3. User Login 35](#_Toc176172304)

[Figure ‎4‑4. Product Search Form 37](#_Toc176172305)

[Figure ‎4‑5. PRODUCT LISTING AND MANAGEMENT 37](#_Toc176172306)

[Figure ‎4‑6. Product Search for Trending 39](#_Toc176172307)

[Figure ‎4‑7. Trend Search History 40](#_Toc176172308)

**Contents**

[Chapter 1 1](#_Toc176562223)

[1. Introduction: 1](#_Toc176562224)

[1.1. Introduction to Digital Platforms and E-Commerce 2](#_Toc176562225)

[1.2. Facebook and its Evolution in E-Commerce: 3](#_Toc176562226)

[1.3. Craigslist as a Traditional Online Classified Platform: 4](#_Toc176562227)

[1.4. Role of Smart Product Hunting Techniques: 6](#_Toc176562228)

[1.5. Significance of Smart Product Hunting in E-Commerce: 7](#_Toc176562229)

[1.6. Purpose and Scope of the Research: 7](#_Toc176562230)

[1.7. Project Timeline: 8](#_Toc176562238)

[Chapter 2 9](#_Toc176562239)

[2. Literature Review: 9](#_Toc176562240)

[2.1. Leveraging Information Retrieval in Customer Product Reviews for Enhanced Product Development: A Multi-Aspect Summarization Approach 9](#_Toc176562241)

[2.2. Sentiment Analysis in E-Commerce Reviews: A Case Study Using Naive Bayes Classifier 11](#_Toc176562242)

[2.3. Advancements and Challenges in E-Commerce Recommender Systems: A Comprehensive Review and Analysis 13](#_Toc176562243)

[2.4. Consumer Buying Behavior Towards Online Shopping in Patna City, Bihar: A Study on Preferences and Influencing Factors 14](#_Toc176562244)

[2.5. Developing an E-Commerce Online Shopping Platform Recommendation Model: Integrating Advanced Algorithms for Enhanced Personalization 14](#_Toc176562245)

[2.6. Fundamentals and Techniques of Social Media Mining: An Overview of Seminal Work 16](#_Toc176562246)

[2.7. Assessing the Impact of Social Media Marketing on Consumer Behavior and Brand Strategy: 16](#_Toc176562247)

[2.8. Current Trends and Challenges in Sentiment Analysis on Social Media: 18](#_Toc176562248)

[2.9. Evaluating E-Commerce Credibility Through Sentiment Analysis: Methodology and Future Directions 18](#_Toc176562249)

[2.10. Understanding User Behavior on Social Media: Stages, Perceptions, and Implications for Business Strategy 19](#_Toc176562250)

[2.11. Analyzing Social Media Interaction and User Interests: 19](#_Toc176562251)

[2.12. Exploring User Behavior and Trust on Craigslist and Online Rental Platforms: Insights from Recent Studies: 20](#_Toc176562252)

[2.13. The Impact of Online Newspaper Advertisements on the Local Economy: Opportunities and Challenges: 21](#_Toc176562253)

[2.14. Advancements in AI-Powered Recommender Systems and Dynamic Pricing Strategies in E-Commerce: 21](#_Toc176562254)

[2.15. Data Mining in E-Commerce: Applications, Techniques, and Challenges in Customer Insights and Social Media Analysis: 22](#_Toc176562255)

[2.16. Opinion Mining and Sentiment Analysis in E-Commerce: Enhancing Consumer Insights and Personalization: 23](#_Toc176562256)

[2.17. The Role of Trust and Privacy in Online Intermediaries: Insights into Feedback Mechanisms and Consumer Confidence: 24](#_Toc176562257)

[2.18. Risk Management Strategies in the Online Marketplace: Identification, Evaluation, and Mitigation: 24](#_Toc176562258)

[2.19. Enhancing E-Commerce with Blockchain Technology: Principles, Benefits, and Research Gaps: 25](#_Toc176562259)

[2.20. AI and Mobile Commerce in E-Commerce: Innovations, Challenges, and Consumer Impact: 25](#_Toc176562260)

[Chapter 3 27](#_Toc176562261)

[3. Methodology: 27](#_Toc176562262)

[3.1. Front-End Development: 27](#_Toc176562263)

[3.1.1. Login and Registration Forms: 28](#_Toc176562264)

[3.1.2. Product Scraping Form: 28](#_Toc176562265)

[3.1.3. Product Listing Table: 28](#_Toc176562266)

[3.1.4. AI-Based Trend Analysis Form: 28](#_Toc176562267)

[3.2. Back-End Development: 28](#_Toc176562268)

[3.2.1. Node.js 28](#_Toc176562269)

[3.2.2. Python and AI Model: 29](#_Toc176562270)

[3.3. Workflow of the Application: 30](#_Toc176562271)

[3.4. Data Scraping Methodology: 31](#_Toc176562272)

[3.5. AI Model Deployment: 31](#_Toc176562273)

[Chapter 4 33](#_Toc176562274)

[4. Results and Discussion: 33](#_Toc176562275)

[4.1. User registration and authentication: 34](#_Toc176562276)

[4.2. Data Scraping Functionality: 35](#_Toc176562277)

[4.3. Product Listing and Management: 37](#_Toc176562278)

[4.4. AI Model Training and Prediction: 38](#_Toc176562279)

[4.5. Deployment of the AI Model: 38](#_Toc176562280)

[4.6. Prediction Accuracy: 38](#_Toc176562281)

[4.7. User Interface Experience: 40](#_Toc176562282)

[Chapter 5 41](#_Toc176562283)

[5. Conclusion: 41](#_Toc176562284)

[References: 43](#_Toc176562285)

**Project links:**

Here are the GitHub and drive links of project resources. All these repositories are public, so anyone with the link can have access to these resources.

**Back-end:** <https://github.com/IbtehajMughal/sph-backend.git>

**Front-end:** <https://github.com/IbtehajMughal/sph-frontend.git>

This link also includes dissertation, presentation and demo video files of project.

**Dataset files:** <https://drive.google.com/drive/folders/1ArF5PU6p9Lz8FzOUgy3xX7As87ZXOfVM?usp=sharing>

# Chapter 1

# Introduction:

This software enables the user to fulfil their product hunting needs by using automation techniques and saves the user time, to enhance their business in a limited interval of time by using just only one click and leveraging artificial intelligence and data mining techniques to improve product hunting. The focus of this project is to develop a software in which user cannot put any type of effort in product hunting as we know that product hunting takes a lot of time and efforts. This project include the development of an AI-based system to predict either a specific product is in trend or not to improve product recommendation based on user search history and interaction with the system and scrap data from Craigslist and Facebook Marketplace which will be discuss in chapter 3 and 4 in detail with all technical aspects.

The ever-changing nature of online mediums has drastically impacted consumers’ approaches to products and services. One of the most important advancements in this field is the emergence of social networks and the Internet as essential e-commerce sites. Facebook and Craigslist can be recognized as key players in this Internet landscape where users are offered specific ways to search, purchase, and use goods. Given that such platforms are now launching their operations and gaining popularity in the market, it has become clear that in order to find relevant products, the use of smart algorithms, data mining, and machine learning is necessary.

Facebook, which started as a social networking site, has transformed into a site through which people do various business-related activities apart from creating, sharing, and interacting with friends and families. The release of the Facebook Marketplace has completely changed the way that users purchase and sell products since it works according to the users’ attitudes, preferences, and friends’ lists. However, due to the large number of products and the nature of users’ interactions on Facebook, it is challenging to find and recommend products to users effectively. At this juncture, we employ strategic approaches to product searching. Through the use of complex algorithms that give insight into the user profile in terms of likes, shares, and search history, these techniques positively impact the product recommendation process, where users will be provided with products that are as close to what they need as possible.

Craigslist, however, stands more or less as a typical classifieds site with a simple interface that automatically excludes a lot of options for modern Internet users, and yet it is perhaps one of the biggest Internet communities focused on the sale and purchase of all types of goods and services. Craigslist, on the other hand, has a minimal graphic interface and is extremely dependent on the users, a condition that may make it very difficult to find certain products. Still, for all its straightforwardness, a loose classifieds format like Craigslist quickly becomes a difficult environment to maintain buyer-seller trust. Product search on Craigslist is one of the key areas that requires the construction of smart systems capable of efficiently and reliably searching through large databases in search of quality and relevant products; other considerations include cases of fraud and the security of users. Ontology-based approaches, natural language processing, and sentiment analysis come in handy in this regard, as these helps provide a robust method of extracting relevant information for improving the efficiency and accuracy of product recommendation systems from text information.

## Introduction to Digital Platforms and E-Commerce

Through the use of digital platforms, customers have interacted differently with brands and products in diverse sectors. ‘They improve the facets of customer relations, organization management, and business ventures in the age of the internet (Tarmizi et al., 2023). This tells us that various factors like brand familiarity, social media marketing, easy accessibility, and word of mouth affect customers’ buying behaviour while purchasing FMCG in online stores (Srivastava and Nalla, 2023). The use of digital platforms also constitutes complementary engagement in digital ecosystems; antecedents as well as specific engagement behaviors and trajectories are illustrated by Davison et al. (2023). A systematic literature on consumer engagement in social networks finds that User and Gratification Theory has been widely adopted in earlier research studies for understanding the subject critically, but this study calls for more work on platforms like Weibo and LinkedIn and in areas that have not been fully explored, such as Africa and Australia (Saikia and Bhattacharjee, 2023).

Social media and e-commerce platforms are currently active components in most business models, especially those of micro, small, and medium enterprises (MSMEs). These platforms present possibilities for markets’ enlargement, improved productivity, and straight-client interaction (Sugiharto, 2024). Although social media may take a lot of time and at the same time provide rather misleading information, it is one of the most effective tools in modern marketing if applied correctly (Fan, 2024; Fitri, Sirait and Batubara, 2024). The findings of Siahaan, Siboro and Saragi (2024) have established the fact that e-commerce adoption has influenced MSME profitability in a positive manner, as has been complemented by social media marketing. However, there are still some issues posing constraints, such as technological issues and digital divides, as well as the risks of electronic attack and information piracy (Fan, 2024; Fitri, Sirait and Batubara, 2024). In order to survive in the environment, digital tools and challenges of the Industrial Revolution 4. Therefore, it is crucial that MSMEs take advantage of these platforms while at the same time remaining adaptive to the implementation of such technologies (Siahaan, Siboro and Saragi, 2024; Sugiharto, 2024). There are some recommendations for overcoming the challenges: stronger government support policies and the development of digital skills for MSMEs’ sustainability in this context (Sugiharto, 2024).

## Facebook and its Evolution in E-Commerce:

Recent work documents the process of the transformation of Facebook from a social networking site to a business entity. Exploring the social commerce attributes integrated by selling platforms, Facebook has become a selling site that supports e-commerce, including the attitudes and buying perceptions of customers (Phan, Đặng and Nguyễn, 2023). Finally, being a website that is available on the World Wide Web, Facebook enables the sharing of content and connections with users from all over the world; it is equipped with different communication means (K, 2024). Due to the existence of competition and the need for young people, the new face of the metaverse was introduced on Facebook. It has featured fresh approaches that can be combined with opportunity zones for property investment (Warner and Sherbia, 2023). In today’s social media world, Facebook is a unique and valuable means of marketing communication that offers coverage and functions that help achieve the goals of selling and communication with the target market (Oktaviani, Maulana and Firmansyah, 2023). Such developments show the processes of turning Facebook into a multi-sided commercial environment from the initial social networking site.

The available literature confirms that the introduction of Facebook Marketplace has shifted buying and selling behaviour, most especially in the areas of customer satisfaction and buying intentions. According to various studies, customer expectation level significantly influences satisfaction, especially when using Facebook Marketplace; other parameters include product quality, delivery, and feedback (Ramos, Arcilla and Macatangay, 2023). Moreover, the findings of the research also suggested that demographic factors have a minimal or no impact on CS. A more recent study on preloved clothing in Kupang, Indonesia, found that shopping lifestyle, social influence, and impulsive buying have a significant influence on purchase intention on the Facebook Marketplace, as indicated by Amaral and Djuang (2023). The following are the implications of these findings for online marketers: Online marketers should endeavour to ensure that they give their customers excellent customer service, deliver their products on time, and that the products do not possess substandard quality, and hence post positive feedback. Furthermore, current lifestyles and social practices should be factored in when businesses are selling used products, as this will boost the number of purchases.

A previous study reveals some of the issues regarding product search and recommendation as the amount of data increases and the variety of users’ touchpoints increases. In study of Liu et al. (2023), the authors proposed a new dataset that consists of real queries, keywords, and full strokes of users’ interactions that overcome the issues mentioned above with simulated datasets. Lafia, Million and Hemphill (2023) identified three types of user navigation paths in research data archives: Life cycle: direct, orienting, and scenic, offering an insight into the strategies and challenges in data acquisition. In their survey published by Quan et al. (2023) discussed big data and artificial intelligence-integrated product design solutions and emphasized that, in order to satisfy various user requirements, there is a possibility to utilize multi-modal data analysis. Van Etten et al. (2023) analyzed data-driven management of crop diversity with a focus on the possibilities for satisfying heterogeneous demands by means of information-dependent management. All these works together emphasize the need for using multiple types of data, analyzing users’ behaviour, and applying sophisticated methods to enhance product search and development in different fields.

## Craigslist as a Traditional Online Classified Platform:

Craigslist is an important player in the online classifieds space, impacting different spheres of internet marketing and marketplaces. Napitupulu, Asti and Hermawan (2023) also present a business model canvas for online classifieds by using artificial intelligence and payment system integration to improve the usability of the platform and increase competitiveness. Sudirman and Disemadi (2023) explained that IPR has issues for Indonesian online marketplaces. Moreover, it is essential to protect IP and promote e-commerce legal frameworks. Saputri et al. (2023) examine how Craigslist-like marketplaces contribute to increasing the competitiveness of Indonesian MSMEs to expand the consumers’ base. Lastly, D’Annunzio and Russo (2021) analyse the ad tech market with a special emphasis on the roles of the intermediaries in the process of information disclosure during the advertising auctions, which is important for understanding the working processes of some platforms, such as Craigslist, in the context of the online advertising environment.

The general idea of simplicity as a language attribute is investigated in the most recent contributions to the field. Simplicity has been realized through proper proportion, omitting anything unnecessary, using durable materials rather than many materials, and the incorporation of green products (Dornubari, Blossom and G, 2023). It results in calming and practical environments that improve people’s quality of life in the building. On the digital front, there are generative agents capable of portraying living persons in interactive settings in ways that are believable: sociality arises from simple starting points (Park et al., 2023). The internet structure has laid down some certain basics that are masked by various complications (Mccauley, Shenker and Varghese, 2023). These studies underline the benefits of a simple design, though it’s significant to point out that the path to achieving simplicity is usually long and implies constant maintenance. The principle of simplicity can therefore be practiced in many disciplines to come up with enhanced practical solutions for people, whether it is designing usable spaces or usable contexts.

There were some credibility and safety issues in video sharing sites as they witnessed problems like fake news and vulgar contents (Niu et al., 2023). Like the nutraceutical industry, challenges affecting the sector include regulatory requirements, efficacy claims, and product authenticity, thus requiring enhancement of systems to enhance sustainability (Komala et al., 2023). In this information technology and digital health sector, the issue of risk management and balancing innovation with saving patient lives then requires the right regulatory measures that have to be promoted differently in different countries, as highlighted by Gilbert et al. (2023). Challenges like compound accessibility, sample size, and bio-genforces are some hurdles associated with natural product-based drug discovery (Simoben et al., 2023). These challenges in the different fields indicate that there is a need to strengthen governance and policies and bring about improvements through implementing solutions that will bolster product safety, effectiveness, and consumer confidence. Solving all these concerns is important for the development and sustenance of these industries as well as the public interest.

## Role of Smart Product Hunting Techniques:

Smart product hunting is the process of using artificial intelligence and big data with the aim of improving the experience of product usage by designing and searching for products. AI can be applied in a number of ways in UX design, including the identification of user needs and requirements, the creation of user experience designs, and even the assessment of the designs (Stige et al., 2023). API-based AI systems in e-commerce will enable the computing real-time data from multiple retailers to enhance search engine results and make the shopping experience more tailored (Journal, 2024). Measurement and analysis of customer preferences, market demands, and product evaluation in diverse data modality and more intelligent product design through big data and AI that bypasses the traditional limitations of design techniques also make their work more efficient (Quan et al., 2023). AI is relevant right from generating the idea of the product to conceptualizing and developing, testing, and even marketing the product. Insightful decision-making and delivery of targeted, individualized marketing, as well as optimal marketing strategies, enhance creativity and competitiveness in product management (Ogundipe, Babatunde and Abaku, 2024).

Social media sentiment analysis is a method of studying the natural language processing for identifying the inclination or the polarity of opinion that is presented on various social media platforms (Ypanaqué and Dueñas, 2023; Aftab et al., 2023). In this task, it uses different techniques based on using vocabularies, machine learning, and deep learning tools to determine the sentiment’s nature as positive, negative, or neutral (Ypanaqué and Dueñas, 2023; Yadav, 2023). In particular, the topic of sentiment analysis has multiple fields of application, namely, the general manifestation of the public’s attitude towards environmental conservation projects, the ability to predict the success of movies, and the consideration of the population’s attitudes during significant events (Yadav, 2023). In the business world, it assists organizations in determining the perception that customers have towards the organization's brand or products (Ahad, 2023). Although sentiment analysis has useful information, problems persist in finding the best models for all languages (Yadav, 2023). As for the ongoing research, it investigates better ways of the technique and also the way of the hybrid methods to increase the classification rate (Ypanaqué and Dueñas, 2023; Aftab et al., 2023).

## Significance of Smart Product Hunting in E-Commerce:

In this contemporary world, e-commerce companies are paying much heed to customers’ experiences, interactions, and loyalty to bring about accomplishments. Some of the most crucial aspects that make customer experiences positive include offers, products, interfaces, and services (Felix and Rembulan, 2023). Personalisation solutions are very instrumental in enhancing the user satisfaction level and increasing the potential of sales by implementing strategies like collaborative filtering and content-based filters, as proposed by Sisodiya (2024). In this case, gamification has been identified as a significant approach in the improvement of customer satisfaction and loyalty to e-commerce firms, and future trends are motivated by the growth of AI and AR/VR technologies (Bogoslov et al., 2023). Perceived hedonic behaviour, system quality, and information quality are the most dominant factors affecting user satisfaction in mobile e-commerce applications, while perceived utilitarian and service quality have comparatively lesser impact (Setiadi et al., 2023). From these conclusions, one can infer insights into getting to know the client, innovative solutions, and patterns that inform and define e-commerce.

## Purpose and Scope of the Research:

The recent studies focus on smart product-seeking solutions on platforms, for example, Facebook Marketplace. One of the models discussed is Que2Engage, which uses embedding-based retrieval to enhance search quality through relevance and engagement (He et al., 2023). Based on multi-task pre-training and hierarchical inference, HierCat, a query categorization system, increases the relevance of search results (He et al., 2023). These innovations have paid big returns in terms of user interaction and site relevance. These include Facebook ads, which have been known to be an effective social media marketing venue for communication and dissemination of smart city programs during the COVID-19 period (Dwianto et al., 2023). In the general eCommerce context, big data and advanced analytics such as machine learning are employed for customers’ clustering, individual approaches, and, of course, price-setting strategies (He et al., 2023). These technologies offer prospects and risks to both the users and the platforms, thereby accentuating the research needed in the field of technological advancement in smart product hunting technologies.

## Project Timeline:



Figure ‎1‑1. Project Timeline

# Chapter 2

# Literature Review:

Technological advancement has played a major role in transforming the facet of e-commerce, including the use of artificial intelligence, machine learning, and data mining techniques. This literature review covers some of the most important topics, such as smart product hunting, which entails the use of intelligent recommendation systems for products and the use of dynamic pricing. It also includes recommender systems in e-commerce, whereby there are basically different techniques that can be applied to improve users’ shopping experiences. The review goes further to look at the social media analysis and mining part, focusing on how the e-commerce strategies and users’ behaviour are shaped with the data from social media. Specific literature on Facebook research and Craigslist research provides understanding of the specifics of users’ engagement as well as trust-related concerns on these Web sites. Furthermore, it analyses the issues of personalisation in the context of e-commerce, including such aspects as recommendation, trust, security, and risk, providing special emphasis to data-driven personalisation and security of the transactions carried out through the Internet. Other advanced strategies, particularly pertinent to e-commerce, include blockchain technology, virtual reality, and augmented reality, which are also discussed as examples for the future developments of the digital marketplace. This review integrates all these aspects to give a comprehensive picture of the current and future of e-commerce.

## Leveraging Information Retrieval in Customer Product Reviews for Enhanced Product Development: A Multi-Aspect Summarization Approach

In a research article by Hou et al., (2019) describe a new use of information retrieval in customers’ product reviews for product development with an emphasis on characteristics other than the physical properties of objects. The proposed summarization model considers five key concepts: Product features, perception of the product, emotional response towards the product, product affordance, and the condition under which the product is used. In this study, the authors therefore intend to use patterns in a natural language to make important recommendations in the design area. The concept development process covers the manual tagging of several online review samples to locate and categorize words and phrases referring to the concepts under analysis. Hence, the cooperation with the other agency analysis of the inter-agreement among human annotators assists in assessing the described linguistic patterns. Significantly, the study focuses on the customers’ ‘contact’ with products, customers’ affective states, and conditions under which products are used to satisfy customers. While discussing the methods existing in the Online review summarization the authors pointed out a few limitations here also including – general characteristic features of the opinionated texts alone, and the unclear specification of data summarizing the customer concerns and engineering characteristics. Such limitations are explained and mitigated in the proposed multi-aspect summarization framework that at least gives the designers an overview of user needs. Using the results of the classification of a set of reviews for the Kindle Paperwhite 3 e-reader, the researchers prove that such a model is useful in the identification of non-feature aspects of design relevant information (Hou et al., 2019).



Figure ‎2‑1 Proposed summarization model (Hou et al., 2019)

### Gaps:

This synthesis will reveal the gaps in current approaches, for example, the neglect of product features in summarisation models, the lack of standardization in definitions in the engineering design field, and the intricacies of language patterns that have not been well addressed.

### Strengths:

It will also demonstrate the benefits of the suggested summarisation model based on the literature review and the consideration of several aspects of users’ preferences, such as product affordance, emotions, and usage conditions. The question of linguistic features and high levels of inter-annotator agreement as the basis for future automated work in the field of summarisation will also be discussed as the key findings of the study.

### Weaknesses:

The synthesis will build on the lacunae of the study, including the use of a small sample size of 265 reviews from a single product. Furthermore, it will also present the issues with a manual process of annotation and the requirement of having a more efficient approach for the summarisation of more extensive data sets.

## Sentiment Analysis in E-Commerce Reviews: A Case Study Using Naive Bayes Classifier

A study by Abighail et al. (2023), the authors present the case of e-commerce reviews and the role and application of SA, specifically from the perspective of NBC which allows for the categorization of reviews into positive and negative. Consumers greatly depend on online reviews in the process of decision-making, thus it’s important to master it. The chosen research sheds light on the issues arising from the ambiguity and the number of reviews and sees the necessity of developing categorization tools. As shown from the previous sentiment analysis on Shopee marketplace data, NBC obtained 72% accuracy and therefore confirmed its efficiency at sentiment classification. This paper focuses on the role of e-commerce in the contemporary world whereby electronic media is used in conducting business through offering and selling products or even services. The advancement in technology especially the World Wide Web has led to the emergence of online social systems that allow users to express their opinions in different aspects of life, be it products or services. In conclusion, the study is important for organizations aiming to improve the level of customer trust and their overall loyalty using the big data of opinionated words and contexts mined from the comments section in the e-commerce platforms.

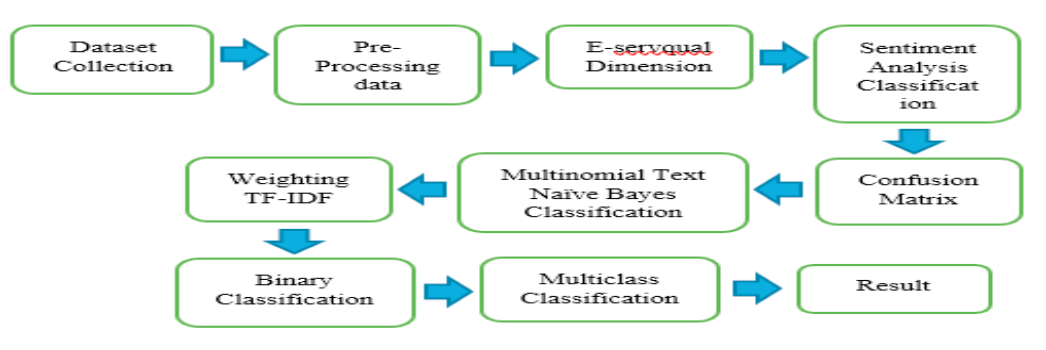


Figure ‎2‑2 Workflow of Research (Abighail et al., 2023)

### Gaps:

Exploration of Emotion Detection: Although the current study aims at the sentiment classification of the reviews, it fails to explore the details of emotion detection within the reviews, which might be of great importance in understanding customers’ feelings and their reasons for writing the reviews.

Impact of E-Servqual Dimensions: While discussing the sentiment analysis results, the link with the e-Servqual dimensions such as trust, reliability, responsiveness, etc. is discussed but not discussed in detail, thus leaving a gap on how the service quality affects the sentiment analysis.

Longitudinal Analysis: The research analyses the reviews of a certain time (the first quarter of the year 2020) and does not consider the changes that may occur in the course of time and can play a crucial role in tracking the dynamics of consumers’ attitudes.

### Strengths:

Methodological Rigour: The study uses the sentiment analysis model, which is based on Naïve Bayes Classifier, a model that is commonly used in text analysis. From table 6, the model has an accuracy of 72%, a recall of 72%, and a precision of 78%, which show that the model is quite effective.

Data Preprocessing: This research proposes a detailed data preprocessing approach that involves normalizing the text data, tokenising the text data, filtering the text data, and stemming the text data to improve the quality of the data used in the analysis.

Relevance to E-commerce: This is very timely and relevant since more and more people are turning to e-commerce. It will therefore be useful to online retailers to gain insights from the current research on how to align their strategies with customer sentiments.

### Weaknesses:

Limited Dataset: This is due to the fact that the data used in the study was collected from 260 reviews only, which can be considered a relatively small sample size. Thus, having a larger dataset would give better results and conclusions.

Single Classification Method: The use of one classification method (Naïve Bayes) may not exploit the possibility of using other classification methods, which may give better results or analysis.

Lack of Comparative Analysis: The document does not specify how the proposed technique compares with other sentiment analysis or classification techniques, and thus the strength or weakness of the approach used cannot be fully ascertained.

## Advancements and Challenges in E-Commerce Recommender Systems: A Comprehensive Review and Analysis

Karimova (2016) stated in his study about the importance and developments of e-Commerce Recommender Systems (RSS) for online settings. This supports the rich personalization and efficient Arch features of RSS which are widely deployed in many online systems. The works published in the last few years are concentrated on overcoming the key limitations of existing RSS, leading to the improvement of their performance elements and reliability. The study also offers a review and classification of over 60 established RS methods for the context of e-commerce and the current state and challenges of e-commerce RSs. In addition, the study acknowledges other directions of research, including the investigation of factors that lead to higher variability in recommendation accuracy such as price. It also raises issues regarding the validity of changing the kind and level of price factors for RSS to enhance competitiveness. The study also points out that the efficiency of RSs plays a considerable role, stating that analyzing every item in a giant e-shop is inadvisable. The experimental results show that there is a way to increase the efficiency of RS while maintaining the relevance of the recommendations provided. In general, the purpose of the study is to present both, directions for practical application and further research concerning e-Commerce Recommender Systems.

## Consumer Buying Behavior Towards Online Shopping in Patna City, Bihar: A Study on Preferences and Influencing Factors

Another study states that, Consumer Buying Behavior Towards Online Shopping in Patna City Bihar India proposed to investigate the consumers’ behavior and their inclination toward online shopping. Data was collected by adopting a convenience sampling technique with fifty responses gathered through Google Form. According to the results, the majority of the respondents showed a preference for online shopping, as opposed to offline shopping, and the majority shopped once a month on average. The most popular products for purchase were electronic gadgets, the second most purchased items were clothes. Consumers thought that shopping online was easy and safe. These factors were based on satisfaction, convenience, price, product quality, and security on the Internet. The study focused on the fact of e-shopping which has gained more popularity among the consumers of Patna City and most of the consumers are moving towards this e-shopping for the fulfillment of their needs. These findings might be helpful for marketers and companies focused on consumers in Bihar, India, and the continuous development of online shopping (Pal and Kumari, 2023).

## Developing an E-Commerce Online Shopping Platform Recommendation Model: Integrating Advanced Algorithms for Enhanced Personalization

The analyzed research article aims to propose an E-Commerce Online Shopping Platform Recommendation Model with synthesized integrated personalized recommendation algorithms to solve the problems of informational overflow and information loss in e-commerce online shopping platforms. As online shopping has become more popular, and personalization is more critical for e-commerce applications, the model uses random forest algorithms, gradient-boosting decision trees, and eXtreme gradient boosting to improve the recommendations’ quality and minimize the sparsity. The research employs bona fide desensitized data from an actual online shopping platform competition, encompassing user behavior data and product usage information. Having removed irrelevant features and optimized the model parameters, the model’s efficiency is examined. Based on the experimental results it can be seen that merging models usually enhances the performance with the fusion of LR, GBDT, and XGBOOST being the best. The possible research directions for the further development of this approach may include real-time recommendations, graphical results, and user-interest model extraction. The model presented below is designed to help a user achieve an improved level of satisfaction with the overall experience in online shopping while being offered relevant and helpful recommendations. Finally, the present study discusses the possible directions of enhancement on the recommendation system and lists down the functional modules of the software system for future research (Xu and Sang, 2022).

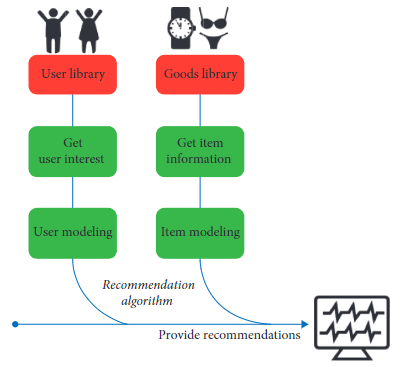


Figure ‎2‑3 Recommendation system model (Xu and Sang, 2022)

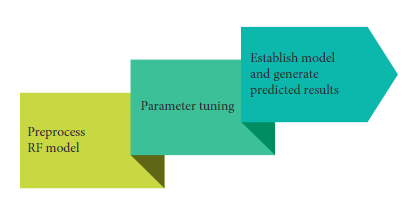


Figure ‎2‑4 Random forest model (Xu and Sang, 2022)

## Fundamentals and Techniques of Social Media Mining: An Overview of Seminal Work

In their seminal work, "Social Media Mining: These fundamentals are outlined by Zafarani, Abbasi, and Liu (2014) in their Book on “An Introduction.” The text’s flow is designed to give readers an understanding of the entire process of social media data mining as a linear process of data collection and cleaning and complex data analysis, considering the features characteristic of social media platforms. The book systematically addresses the key areas of knowledge including the basics of data mining and social network analysis and provides detailed information on the peculiarities of peculiarities of the problems and approaches used in large-scale social media data analysis. It provides several algorithms and techniques that are useful for such tasks as community detection, sentiment analysis, and information diffusion. The authors also mention the issues of ethical concern and violation of privacy that are more significant in the case of social media mining. Real-world examples and cases are also incorporated profusely into the text to make the pedagogy more applicative for its audience. Moreover, the book also covers the topics related to the tools and programming languages that are important for practicing social media mining namely Python, and the libraries used along with it. Overall, "Social Media Mining: To the target readership base of both scholarly and business-oriented researchers and practitioners, ‘An Introduction’ is quite useful. Not only does it give a clear theoretical approach, but it also delivers an outline of the tools and methods for the implementation of efficient data extraction from social networks. These are useful for a researcher who wants to study and work on social media mining where he or she is guaranteed to have the right knowledge as well as apply proper ethical measures while doing the work.

## Assessing the Impact of Social Media Marketing on Consumer Behavior and Brand Strategy:

The viability of social media marketing as a marketing strategy and its impact on persuading consumers. It underlines the importance of the concept of innovativeness in the acquisition of more market niches using template networking. Overall, the study seeks to establish the extent to which social media influences a firm’s reputation, branding strategies, and product promotion. It emphasizes the importance of monitoring consumer interactions on social network sites for enhancing the product and adjusting the marketing plan. Furthermore, the research entails an analysis of how global sales and @SamsungMobile Twitter follower volumes have changed over five years. The study proves dynamic change in consumers’ behavioral patterns and implies directions for improvement and diversification of like brand research (Singh and Singh, 2018).

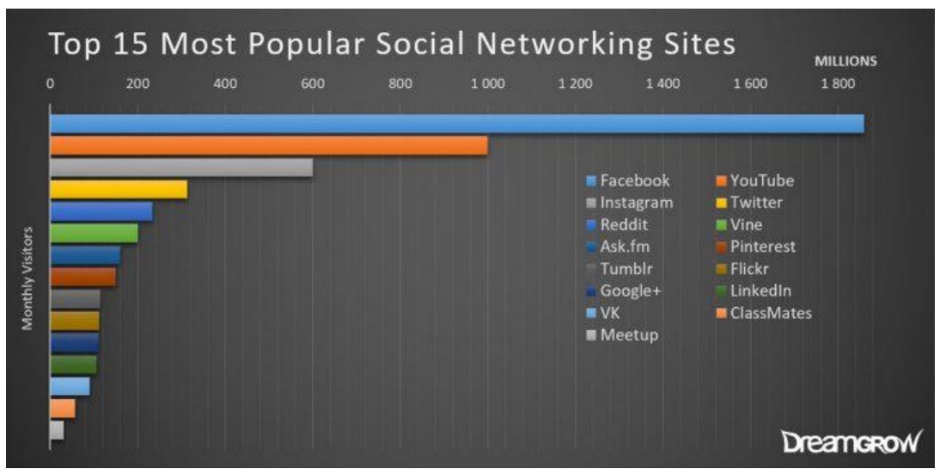


Figure ‎2‑5 Most Popular Social Networking Sites (Singh and Singh, 2018)

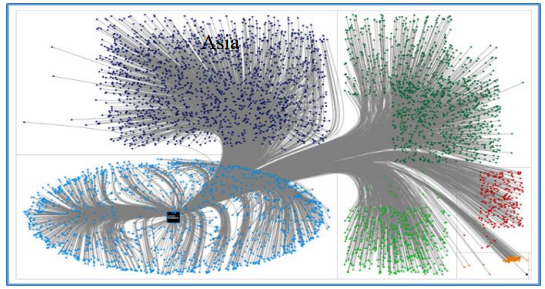


Figure ‎2‑6 @Samsung Mobile Followers across (Singh and Singh, 2018)

The global ranking of the Top 15 social networking sites is depicted in Figure 5. The figure presented is a bar chart showing the ranking up to February 2017. Such platforms include Facebook, which tops the list, followed by other important platforms such as YouTube, Instagram, and Twitter. The circulatory shape symbolizes the continually increasing reliance on social media, which is important in the marketing of businesses. Furthermore, it gives an understanding of the competitive environment of social networking sites at that time and demonstrates their importance in the sphere of digital marketing. Figure 6 shows relationships among 5,000 actives @Samsung Mobile followers across continents, with a focus on the connection. The graph helps to constantly illustrate where in the geography of the world the shared info is mostly spread concerning the Samsung followers. This distribution demonstrates the fields that are currently the most active, thus reflecting the worldwide connections of Samsung and the impact of social media.

## Current Trends and Challenges in Sentiment Analysis on Social Media:

The current trends in sentiment analysis on social media have involved the use of such models as CNN-LSTM and Stacked-BiLSTM that have been tested on movie reviews with good outcomes. Also, deep learning models have been effectively applied to analyzing sentiment in the tourism and restaurant sectors with high precision. Nevertheless, the following challenges exist; Data quality issues, the problem of dealing with a huge amount of data when classifying, and the need for efficient methods of processing data particularly in real-time applications. Meeting these challenges while adopting the current trends will be vital for the future development of sentiment analysis on social media (Xu, Chang and Jayne, 2022).

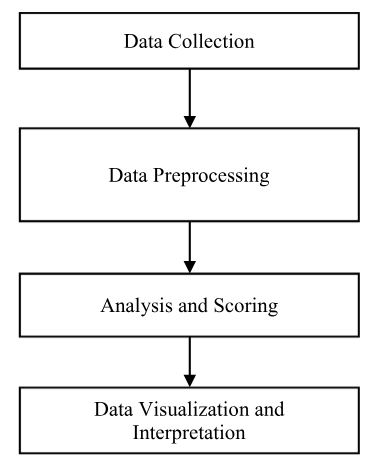


Figure ‎2‑7 The general process of sentiment analysis (Xu, Chang and Jayne, 2022)

## Evaluating E-Commerce Credibility Through Sentiment Analysis: Methodology and Future Directions

It is important to assess the credibility and trustworthiness of e-commerce businesses and social media is an effective tool to do this, especially through the use of sentiment analysis. For the classification of data in this study, six classifiers were used, where Logistic Regression gave the highest accuracy of 76.7%. The analyses from comments on social media assist consumers in identifying the reliability of a product and the specific benefits it offers in the long run. Possible future works include improving the proposed model, removing the spam in the comments, and capturing the sentiment expressed through emojis to extend the application of sentiment analysis in the e-commerce context (S et al., 2023).

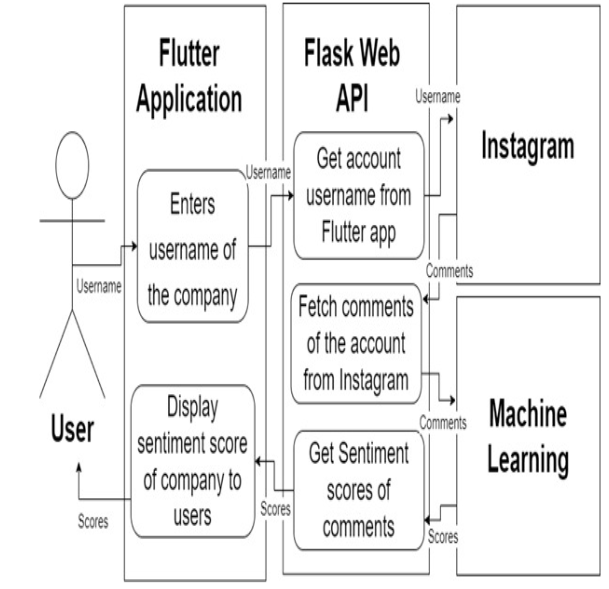


Figure ‎2‑8 Methodology used (S et al., 2023)

## Understanding User Behavior on Social Media: Stages, Perceptions, and Implications for Business Strategy

To study the behavior of users on social media, there are several stages involved such as gathering data from different sites, cleaning the text data, and classifying user behavior by organizational needs. The purpose of this study is to identify users’ perceptions, attitudes, and emotions towards various contents that are posted on social media platforms about products, services, organizations, persons, events, and topics. This is because the behavior depicted by the users can be used to make the right decisions, provide the right strategies, and design appropriate products and or services that can suit the needs of the users in the market and hence improve the performance of the business in the digital marketplace D. R. (2018).

## Analyzing Social Media Interaction and User Interests:

A study by Viswanath et al. (2009) analyzes user interaction on Facebook. It examines over 60,000 users and over 800,000 logged interactions over a period of two years. The study investigates both the microscopic level of interactions between user pairs and the macroscopic level of how these interactions affect the overall structure of the activity network. The findings suggest that the activity network evolves rapidly, with many interactions being influenced by site mechanisms like Facebook’s birthday reminder feature. In the same context a research paper by M et al. (2021) includes methodologies and algorithms for predicting user interests on social media platforms. It discusses the use of supervised machine learning approaches to categorize users based on features such as tweet content, user behavior, and time-series data. These methods can be applied to platforms like Facebook to predict user attributes, preferences, and interests, enhancing the accuracy and scalability of user profiling algorithms. A research paper by Arthur (2023) some of the data privacy issues affecting e-commerce businesses are centered on acquiring appropriate consent, data security, and conformity with data protection laws. Customers are usually wary of over complicated consent procedures, resulting in misconceptions and detrimental effects on a firm’s reputation. Hence, firms need to declare accurate privacy policies, enhance protection, and seek advice from experts on the legal requirements. It is important to address these challenges optimally to retain customer trust as well as preserve individuals’ information.

## Exploring User Behavior and Trust on Craigslist and Online Rental Platforms: Insights from Recent Studies:

Studying user behaviors on Craigslist means analyzing the dynamic processes of users’ activity about a variety of categories on the site, available for jobs, housing, goods, etc. A survey carried out by Park et al. (2017) indicated that the users’ search inputs leaned more toward localities, hence the significance of geographical context in dynamics. Moreover, it means that due to the site’s simplicity and anonymity of communication through Craigslist, the majority of users positively define the level of activity on this platform (Chen et al., 2021). Other research also reveals that user trust plays various aspects such as reliability of listing or other user interactions, community standards, and a reputation system (Goddu & Goddu, 2024). A research paper by (Kennedy, 2023) stated that security is also an essential factor when it comes to online transactions, especially in the case of renting premises for holidays since such a type of service has already been touched by fraudsters. It is necessary to defend buyers and sellers with the help of careful merchandise inspection, payment safeguarding, and conflict resolution. To these factors, KOALA responds by holding the money until check-in to guarantee safety for both the providers and consumers and popularizing the market. Eliminating blurred images of the members’ profiles and genuine customer testimonials adds to the overall level of trust and believability of the site, making it safer and even more attractive. In this way, timeshare rentals are made more convenient for the users, risks to the vacationers are minimized, and barriers to the rental process are eliminated for the owners and the customers of KOALA.

## The Impact of Online Newspaper Advertisements on the Local Economy: Opportunities and Challenges:

Newspaper advertisements have moved from the print medium to the Internet which has affected the local economy. Selling advertisements online improves a business locally, and provides employment opportunities and inexpensive advertisement platforms. But they also compete with conventional business formats such as local papers. These platforms expand market access, cause lower advertisement prices, and enhance control. Thus, despite certain imperfections, for instance, frauds and spam, classified ads websites remain to unite people and companies and influence the way we work and communicate (Khan, 2023).

## Advancements in AI-Powered Recommender Systems and Dynamic Pricing Strategies in E-Commerce:

A research paper by Zhang et al. (2020) focuses on the practical aspects of AI in the recommender systems that are used for recommending certain products to customers. Hence, the paper identifies the kinds of methodologies and techniques that are used in recommender systems, and how the incorporation of AI: Computational intelligence, and Machine learning techniques leads to better accuracy of prediction and tackles issues such as sparsity of data and cold start problems. It also outlines existing research questions and directions for a future study focusing on the reinforcements of recommended systems by AI techniques including fuzzy schemes, transfer learning, genetic and evolutionary algorithms, noise-tolerant neural networks, deep learning, and active learning. Also, nowadays dynamic pricing is considered one of the most effective strategic weapons in e-commerce that involves price changes frequently according to demand, competition, and various other factors. This method enables firms to maximize their revenues by using rather complex algorithms and data processing. Kinds include peak pricing, customer-specific pricing, group classification pricing, and temporal pricing. Some of the well-known companies, such as Uber, Amazon, and Airbnb, can be mentioned as using dynamic pricing strategies, which allow them to adapt their prices to the position of the supply and demand and customers’ actions. The key activities that define the process of dynamic pricing consist of the following ones: data collection, the formulation of strategy, decision on the proper tools, pilot implementation, and monitoring. Although it increases revenues and customer involvement, its use is sensitive when carried out inappropriately as it can lead to customer dissatisfaction and negative consequences for a company’s image (Okhrem, 2024).

## Data Mining in E-Commerce: Applications, Techniques, and Challenges in Customer Insights and Social Media Analysis:

Data mining has been defined as a new focal research direction and activities, which are carried out by statisticians, artificial intelligence researchers, and database engineers because this field can reveal numerous patterns that may strengthen the forecasts and personalized approaches in e-commerce environments. The uses of data mining in the e-commerce domain are primarily focused on enhancing changes that produce value to the customers including identifying new customers, maintaining current customers, and estimating customer purchasing behavior. Key application areas include customer profiling where data mining plays a critical role in the process of analysis to bring out customers’ behavior and preferences and to ensure that products and services are developed to suit the customers’ preferences. The various practical examples elaborate on the necessity of aggregating the data at the right granularity level and opine that gaining insights by session Web logs or leveraging data from the application server is more valuable for understanding customers’ transactions (“DATA MINING AND E-COMMERCE: METHODS, APPLICATIONS, AND CHALLENGES,” 2008).

Data mining techniques are essential for the analysis of social media because of the massive and constant streaming of real-time data on these sites. In this case, the more traditional forms of statistical analysis prove to be inadequate in dealing with this large volume of data. Some of the common methods include AdaBoost, along with sentiment analysis, and opinion mining. However, there seem to be no sufficient comparative studies concerning the accuracy and performance of these techniques as pointed out and marked by literature review (Injadat et al., 2016). A similar study presents a general and more encompassing classification to enlighten the general scope of BDA and the contribution of the latter in developing business value within the e-commerce context. First, the paper outlines the definitional perspectives on BDA and defines characteristics and types of big data within e-commerce Second, it shows how big data delivers business value in the context of e-commerce Third, the paper offers suggestions for overcoming the challenges concerning the application of big data in e-commerce (Akter & Wamba, 2016).

## Opinion Mining and Sentiment Analysis in E-Commerce: Enhancing Consumer Insights and Personalization:

Opinion mining and sentiment analysis are the branches of computational linguistics that focus on using technologies to understand people’s opinions, attitudes, and feelings toward entities, individuals, issues, events, topics, and their characteristics. This field I have found can be quite technical but, when applied it has real-life benefits, especially to businesses conducting polls on the consumers and customers on the other end searching for feedback on the products. However, thanks to the emerging flood of social media sites, the task of looking for, screening, and then evaluating and summing up opinions has added a new heaviness because of the final complexity of text that contains opinions of different persons. Automatic approaches are required to eliminate subjection, and mental restrictive aspects that can be observed in manual analysis (Liu & Zhang, 2012). Sentiment analysis in e-business targets one of the significant problems of consumers, namely, their inability to touch a product or service before obtaining it. Sentiment analysis makes it easier to categorize each review as positive, negative, or neutral from the summary produced from the review depending on the features of the product. This in turn helps consumers, owners of e-commerce stores, and manufacturers acquire up-to-date and easily understandable information on their customers and make better decisions when undertaking online shopping (Basani et al., 2019). Customization in e-commerce is crucial for improving the UX as it delivers the needed services to the consumer with comfort. Personalization creates a better user experience by analyzing data and applying specific changes to the user’s shopping process. This can be of great help to any e-commerce store since clients will be able to find products that suit them without much struggle (Vavliakis et al., 2019).

A concept that is comprehensively explained by (Raorane & Kulkarni, 2011) is data mining, with extended reference to the process as a tool for evaluating consumer behavior with special regard to supermarkets. Analyzing the patterns of spending time, choice of time for shopping, and choice of products that frequently hit the consumer markets would go a long way in helping data mining to better position strategies relating to business ventures. Thus, association rules and other algorithms are used to extract useful rules and patterns in an attempt to enhance the knowledge concerning customer preferences and actions.

## The Role of Trust and Privacy in Online Intermediaries: Insights into Feedback Mechanisms and Consumer Confidence:

The concept of trust plays a significant role for every online intermediary like eBay Amazon Uber and Airbnb in the world. Feedback is established by the trust metrics mainly associated with reviews that enable buyers and sellers to rate each other, as well as the transacted products or services. They assist in determining who amongst the participants gets a certificate of being trustworthy, encourage participants to engage in the right way, and assist in contracting between strangers as well as offering vital information about the stranger (Luca, 2017). Privacy issues in online purchasing involve acquisition, management, and precaution concerning the privacy measures implemented. Web users’ perceptions related to the manner in which their personal information is being gathered, processed, and protected by websites directly impact upon their levels of trust and perceived risk in e-commerce settings. Nevertheless, the measure of consumers’ awareness regarding their activities on the web is increasing, and that contributes to the confidence in electronic purchases as well (Gurung & Raja, 2016).

## Risk Management Strategies in the Online Marketplace: Identification, Evaluation, and Mitigation:

The risk management process in the online marketplace aims at – Risk identification, risk evaluation, and risk mitigation. The strategies involved are; Risk minimization which is by avoiding doing business with unscrupulous suppliers or shifting from credit sales to merchants that tend to default in payment, increasing security status and exiting high-risk geographical locales, and Risk distribution through negotiating contracts with other partners in the value chain. Besides, conflict and privacy issues are vital for financial risk management even though they may seem general (Miranda, 2022).

## Enhancing E-Commerce with Blockchain Technology: Principles, Benefits, and Research Gaps:

In particular, the principles of blockchain technology are applicable to raise the usage of e-commerce by delivering transaction safety, speed, and clarity. It makes it possible to develop peer-to-peer networks that will enable users to securely store and transfer their digital resources to minimize cases of fraud and leakage of information. Nevertheless, the subject of realization of e-commerce with the help of blockchain is still insufficiently investigated and more investigations are still needed to solve a great number of challenges associated with that field (Taherdoost & Madanchian, 2023).

## AI and Mobile Commerce in E-Commerce: Innovations, Challenges, and Consumer Impact:

AI in e-commerce has become the new standard as it has changed how product recommendations, conversational interfaces, chatbots, and forecasting and prediction are done. However, its strengths should come hand in hand with other problems and drawbacks, like data quality errors, algorithm biases, integration issues and questions of ethics. Innovations such as hyper-personalization, augmented reality, and the proper implementation of ethical AI give a glimpse of more valuable advancements to come in the sphere of digital marketplace (Kaur et al., 2024). Mobile commerce commonly referred to as m-commerce was discussed in a paper by Ettis and Abidine (2017) with a specific focus on aspects affecting the consumer. Finally, the factors that affect m-commerce are as follows: Consumer characteristics, perceived usefulness, perceived ease of use and perceived communication odium, and social influence. The research further urges the need to consider these variables in fashioning the m-commerce plans to increase consumer satisfaction thereby improving the chances of customer loyalty.

Many elements of e-commerce are nowadays influenced by versatile technologies, including AI, machine learning, and data mining, which improve product recommendations provided to users, the level of their engagement, and protection. Social media importance is still unbeaten when it comes to analysing consumer behaviour and business marketing techniques. Other related technologies, such as blockchain and augmented reality, one can expect still greater improvements. However, there are concerns like data privacy, bias in algorithms, and arising consumer needs that have to be looked into for these technologies to reap maximum benefits. If these set challenges are met, then the e-commerce sector will be able to transform the current marketplace into a highly personal, secure, and efficient digital landscape.

# Chapter 3

# Methodology:

This chapter gives an outline of the method employed in the development and application of the aforementioned proposed application, aimed at enhancing product scraping and capability in trend anticipation through the employment of artificial intelligence (AI). This particular application synchronizes several technologies as a way of enhancing its functionality successfully both at the front and the back end in order to achieve its goals and objectives.

The front end of the application is developed in React, which is an efficient and versatile JavaScript library for the development of the front-end. The front end consists of various forms and web elements associated with it, like the user registration process form, login form, product scraping form, and artificial intelligence-based trending update analysis. Such components allow the users to enter all the needs, work with scraped product information, and request predictions from the AI.

The back end is supported by a mix of Node JS and MongoDB to provide easy navigation to visitors in between site pages. js where it is most applicable and Python where it shines. Node.js has its means of handling user authentication, product scraping, and data management using API endpoints through its Express server framework. Python is used for the AI part, where a machine learning algorithm, the Random Forest Classifier, is used to predict data trends from past product data. This model is run using Flask—a lightweight Python web framework—to develop a real-time prediction dashboard.

The methodology guarantees the organismic approach to the technologies’ integration and thus provides the application that offers a solid and intuitive environment for product analysis and decision-making.

## Front-End Development:

The part of the application that will be presenting the user interface has the contacts application front-end designed with React JavaScript framework. The front-end comprises several components, including:

### Login and Registration Forms:

One of them is that a single user starts the process and registers through a form. After completing the registration process, the user will be directed to the login page after registration is done. When logged in, a user will be able to use different functionalities of the application.

### Product Scraping Form:

Upon signing in, the users are immediately directed to the product scraping tab where they can fill in details like the product name, the price range that they wish they saw, and the platform from which they want to scrape data, whether Facebook marketplace or Craigslist, for instance.

### Product Listing Table:

This table shows all the previously scrapped products one under the other, enabling the user to view the scrapped data in an orderly fashion.

### AI-Based Trend Analysis Form:

There is a separate form where a user inputs the link to an Amazon product. The AI module embedded into the application recognizes the product link and forecasts the popularity of the product within the recent period.

## Back-End Development:

The back-end development is a blend of Node, Express, and MongoDB, which are progressive platforms and an indication of growth for this project. JavaScript and Python for different processes related to perform AI model tasks on this website.

### Node.js

Node.js for the server-side, the JavaScript runtime based on Chrome’s V8 JavaScript engine, is used for managing the server part. The Express server is implemented, and the server node is running on the local host. JS to define multiple routes and handle API requests. The server includes routes for:

User Registration and Login: When a user registers for a new account or is using a previous account that he or she has been using. It deploys an API for authentication and supervising a user’s session to make it safe.

Product Scraping: When a user asks to scrape a given product from Craigslist, an API call is sent to the Node.js server. The server then uses a package known as "Puppeteer," a headless browser, to scrape data from Craigslist concerning the product in question.

Product Listing: Like the case with product scraping, the server also handles the listing of products via routes, which utilize the database and display previously scraped products.

AI Prediction Model: A third route is formed to access the AI prediction model, which is running on a Python server. It will give the front-end the opportunity to send analytics to the AI model and, in return, get the results of the prediction.

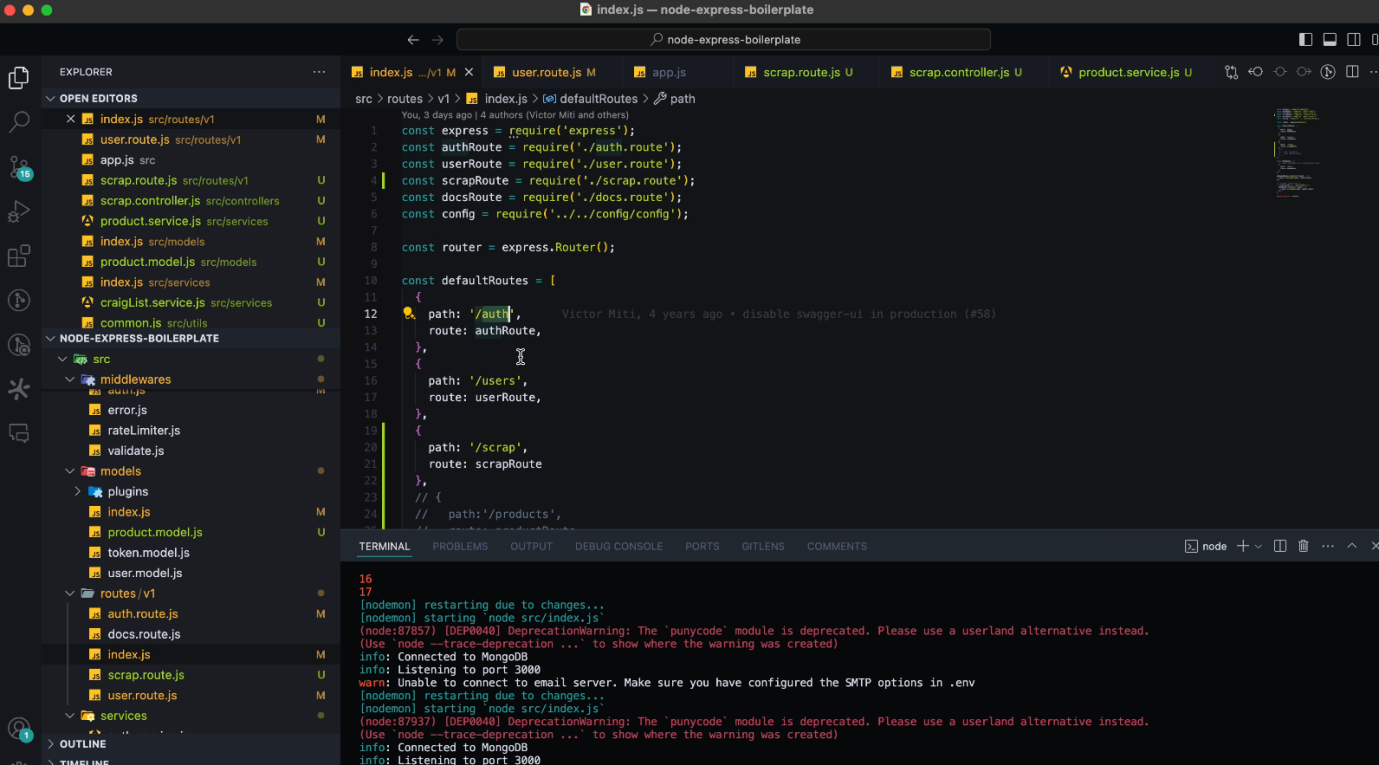


Figure ‎3‑1.Routes

### Python and AI Model:

Python is used to construct and train the AI model for the forecast of product trends. The Math and ‘numpy’ libraries are used to manipulate data, while Scikit-Learn or'sklearn’ is used for applying machine learning techniques. The selected algorithm for AI forecasting is the Random Forest Classifier. The development process of the AI model is as follows:

Data Collection and Preparation: Information is gathered for a particular time, for instance, ratings of products,product name, number of units sold in previous month, customers’ feedback, and other appropriate statistic indices. The dataset is split into two parts: This is because typically the data is split in the ratio of 8:2 for training and testing, respectively.

Model Training: This algorithm that is used to train the model is known as the Random Forest Classifier. One of the steps of the training process is based on historical data to discover such patterns to find out whether this product is trending or not based on product rating and number of units sold.

Model Evaluation: The testing dataset is then used to assess the model to check on the accuracy and performance of the model. The final trained model of the text data is stored and saved for real-time purposes using the Pickle format and the file extension PKL.

Integration with Back-end: The trained model is then incorporated into a Python server through the Flask library. Using the “predict” route to load the model, process the data, and provide the results in the form of an array of 0 for not trending and 1 for trending.

## Workflow of the Application:

User Registration and Login: A new user signs up from the front-end, and this process invokes an API that will be hosted on the Node.js server. It identifies the user via some identification such as email or user login, password and grants the user entry into the application.

Data Input and Scraping: The user starts with entering the product details on the scraping form. This particular application makes a call to the node, which gets it as a request. Node.js server that invokes the Puppeteer library to grab information from Craigslist or any other required website.

AI-Based Trend Prediction: In case of trend analysis, the user enters a product link for Amazon. This link is sent by the application to the Flask server, and the AI model, which is in Python, investigates the data and provides a trend forecast.

Displaying Results: In returning, the back-end sends the result to the front-end to show the user whether or not the product is popular.

Product Management: Users are also able to control their scraped products via the product listing, which shows data fetched from the back-end database in a structured format.

## Data Scraping Methodology:

The data scraping process forms a part of this application to help in the efficient and accurate scraping of product details from external sites such as Craigslist and Facebook marketplace. The scraping process is performed by the puppeteer package, which is a giant tool in the Node environment. Node.js environment for headless browser automation scripting and the to, respectively automatically. Puppeteer is in fact the headless version of the Google Chrome browser, with the ability to interact with web pages, imitate the user actions, and even scrape data without having to use the graphical interface. This makes Puppeteer especially useful for large-scale data extraction scenarios since it doesn’t have to work in a full browser environment, which renders it both faster and more economical in terms of the resources it consumes. In this application, Puppeteer is utilized to carry out the process of web scraping by extracting product data like names, descriptions, price, and other attributes from Craigslist and Facebook marketplace. The scraping initiation process starts when the user enters a request through the front-end application, which features the target platform and product attributes. Then an API call is made to the node for sending messages to register with the event emitter. refers to the JS server, which in turn launches a headless browser with the help of Puppeteer and performs search operations. Puppeteer moves to the specific URLs of the Craigslist web site and Facebook site, scraps/collects the following data, and collects them while being compliant with the data access policy of those two aforementioned web sites. This adherence is important due to restrictions in terms of service of such platforms, which may include rate limiting challenges, and data usage. Furthermore, the application addresses such concerns seamlessly through leveraging the solution’s automation provided by Puppeteer, using different user agents. The scraped data is next cleaned, normalized, and stored in the application’s database, where they are in a ready and accessible state to be utilized for processing and presentation. Combined with the sets of features offered by Puppeteer, the application can achieve consistently high rates of scraping precision, which will be beneficial when providing users with timely and pertinent product data for decision-making and market trend evaluation.

## AI Model Deployment:

The further deployment of the AI model is realized using the Flask server, which is regarded as the back-end to host the ML model as well as to create a RESTful API interface. This input entry point is developed with the specific purpose of quickly receiving inputs of products from the front end of the application, so as to avoid complicated information exchange between the front end of the application that is accessible by the end user and the AI prediction algorithm. When product data is submitted, the Flask server parses it and passes this data to the specifically trained Random Forest model, which was built with the help of the Pandas library for data handling and the Scikit-Learn library for model training. Using the information relating to product-related data from the past, the model rests on the ability to predict the pattern of change or even the behaviour as to whether a certain product is currently trending or not. After it has received this data, the AI model immediately runs its algorithms so as to produce a prediction result and then send this to the front-end in real time. The kind of output given by the prediction model is in discrete form, yes or no, or in terms of a binary form: 0 or 1 represent, respectively, no trend towards the product and a trend towards it. Such feedback can be offered instantly and can be of great help in decision-making while also improving the general efficiency and quick response to various actions of the application. The Flask-based deployment ensures that the deployed model is fully functional and able to handle a number of requests simultaneously and, at the same time, sets the bar on performance high enough.

To develop a more engaging interface together with an AI model for predicting trends, this approach integrates both back-end and front-end technologies. The combination of React for front-end development, Node.js for back-end operations, and Python for learning, as well as deploying the AI model, results in a strong, scale able application structure. For instance, it allows efficient data extraction, administration, and immediate forecasting, thus making it an important resource in analyzing patterns in product development.

# Chapter 4

# Results and Discussion:

The ideas concerning product scraping and trend prediction, together with the application to be developed, have been successfully implemented with the help of both front-end and back-end tools. The front-end of the application was created with the help of React for responsiveness, which includes basic forms like log-in, registration, product scraping, and the form that shows a list of products. On the back-end, Node.js and Python were used for many operations, such as user authentication, web scraping, and AI-based trend analysis.

The practical applicability of the application was determined through several steps. First, both the registration form and the login page were checked for efficient user engagement. On the successful registration, the users are taken to the login page, after which the authentication is presented with a form that needs the user to input the product name, the price range, and the platform to be scrapped. The functionality to scrape product data, including description, price, available quantity, etc., from Facebook marketplace and Craigslist was validated using Puppeteer, a headless browser module in Node.js. The data scraping exercises were performed with great success, and the effectiveness was highlighted in the ways the identified application is able to adhere to the data access policies on the intended platforms with very high levels of data scraping efficiency.

The feasibility and practicality of the integration of the AI model for trend prediction were done on the back end using Python. To train the model, I used 80% of the data set, while the remaining 20% was used for testing the trained model; meanwhile, the data manipulation and reading were done in the Python environment using Pandas. The model was checked for the viability of forecasting whether a product is likely to trend or not based on its past interactions. The fine-tuned model was saved as a PKL file format so as to enable it to process the real-time data in real time. The service function of Flask provided the hosting of the model and the establishment of a RESTful API point, namely “predict,” which allows the product data to be processed and returned prediction outcomes in real-time when received from the front-end.

The results found here show that the proposed AI model has been reasonably stable and acceptable in its predictions of product trends, with the output being of binary nature, 0 or 1, implying that a specific product is or isn’t trending. This output was easily embedded on the user interface of the application, whereby we were able to give an immediate response to the user. The prediction accuracy from the testing set shows the effectiveness of the Random Forest algorithm in product trend analysis and when complemented with fundamental pre-processing and data manipulation.

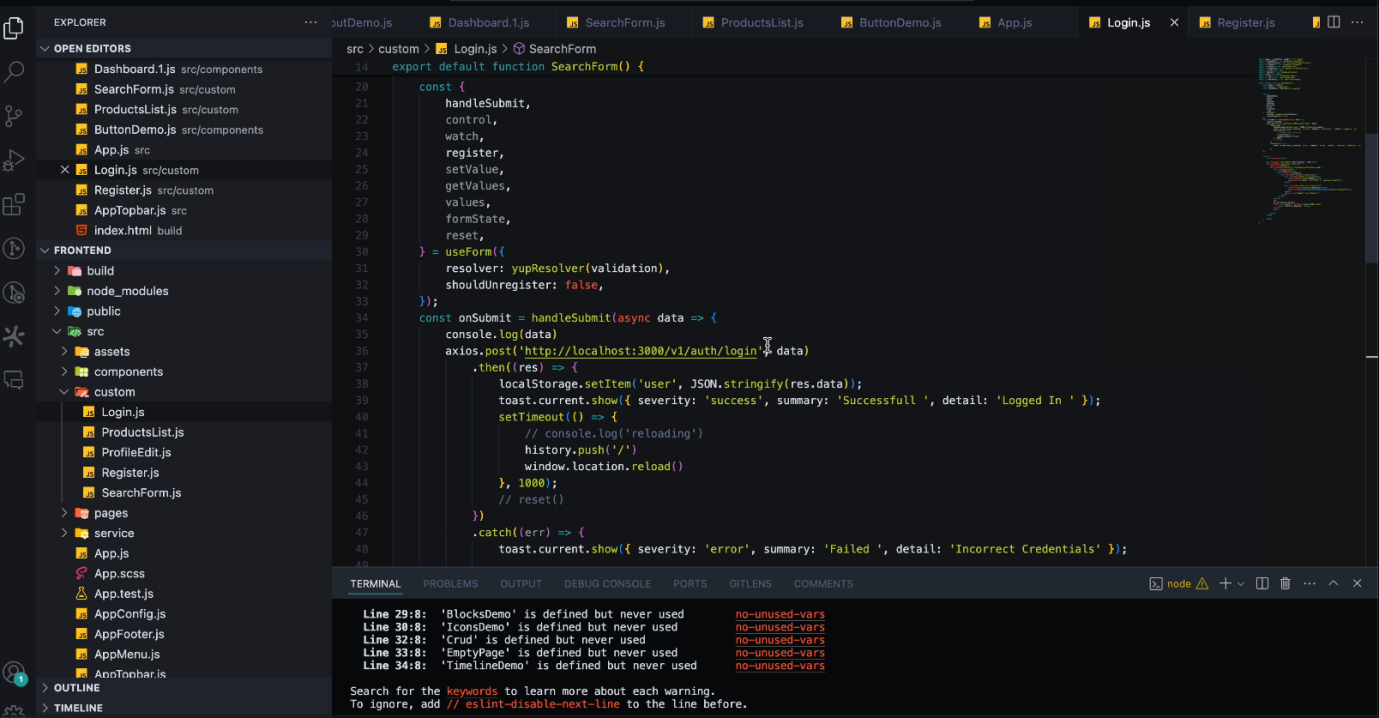


Figure ‎4‑1. Back-end Structure

## User registration and authentication:

The use of the application was able to include the ability to register users as well as the ability to authenticate them. Students could sing up, and after that, if they filled out the out the form on sing up, they were taken to the log-in page. The back-end was efficient in managing user information and checking for registration and credential validation during logins for secure and efficient users’ access.

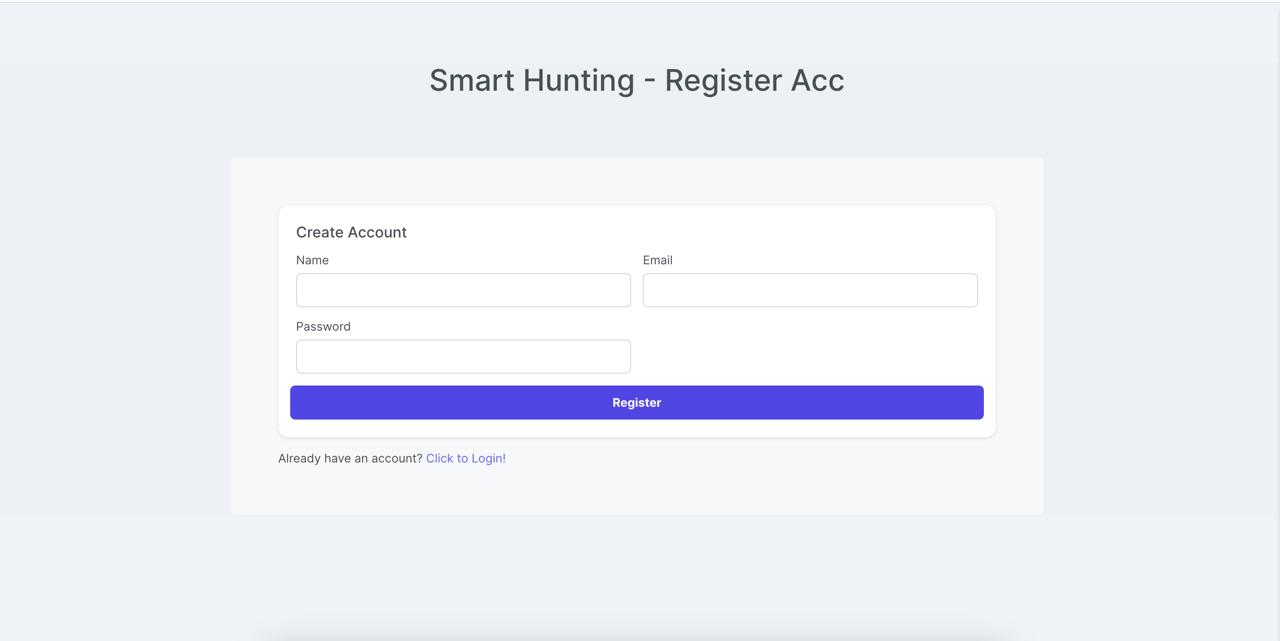


Figure ‎4‑2. USER REGISTRATION AND AUTHENTICATION

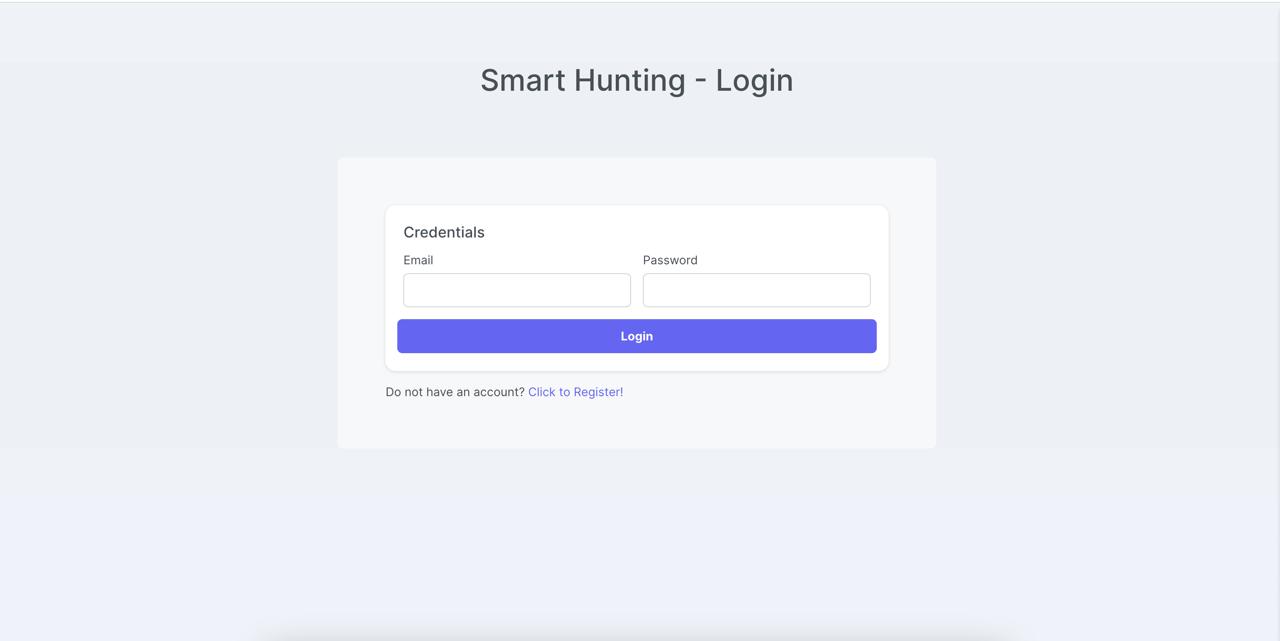


Figure ‎4‑3. User Login

## Data Scraping Functionality:

The data scraping functionality was then tested, whereby the application was used to scrape for product information from both Facebook marketplace and Craigslist. Using the Puppeteer package in Node, in order to automate the filling of form data.js, the application was able to get through the target web pages, simulate user interactions, and collect the required product information such as names, descriptions, and/or prices. All of the scraping was done following the data access policies of the specific platforms, and the scraping methodology outcomes showed high accuracy and reliability.

Various problems were faced in the product scraping section, which had to be addressed to make the project a success. The first major issue was the issue of scraping efficiency, especially regarding the dynamic content that was prevalent on sites like Facebook. Due to the fact that data on Facebook marketplace is loaded dynamically, the puppet package was set to wait for certain periods of time and to ‘listen’ for certain page events to ensure that all necessary elements were loaded before scraping. This was useful in mitigating delays and made quite certain that the information obtained was valid and reliable.

There was also the issue of platform boundaries and constraints, including rate limiting and bot protection mechanisms. Such restrictions meant that the scraping requests had to be rationed so as to avoid detection and subsequent blocking by the platforms. The solution was to make 5 seconds of delay between actions and make the program work like a real human when submitting a form and navigating through the page. Also, the issue of data access policies of the platforms was well observed in order to make sure that the scraping activity was within the legal and ethical frameworks.

Data extraction from different platforms was handled by the difference in page layouts and formats. To address this issue, the scraping logic was made to be intelligent, using CSS selectors and XPath queries to learn about the location of the data elements on the web page, thus scraping. The system was fine-tuned to catch product names, descriptions, and prices with high precision for both the Facebook marketplace and Craigslist platforms; this is proof of the technical effort put into the project.

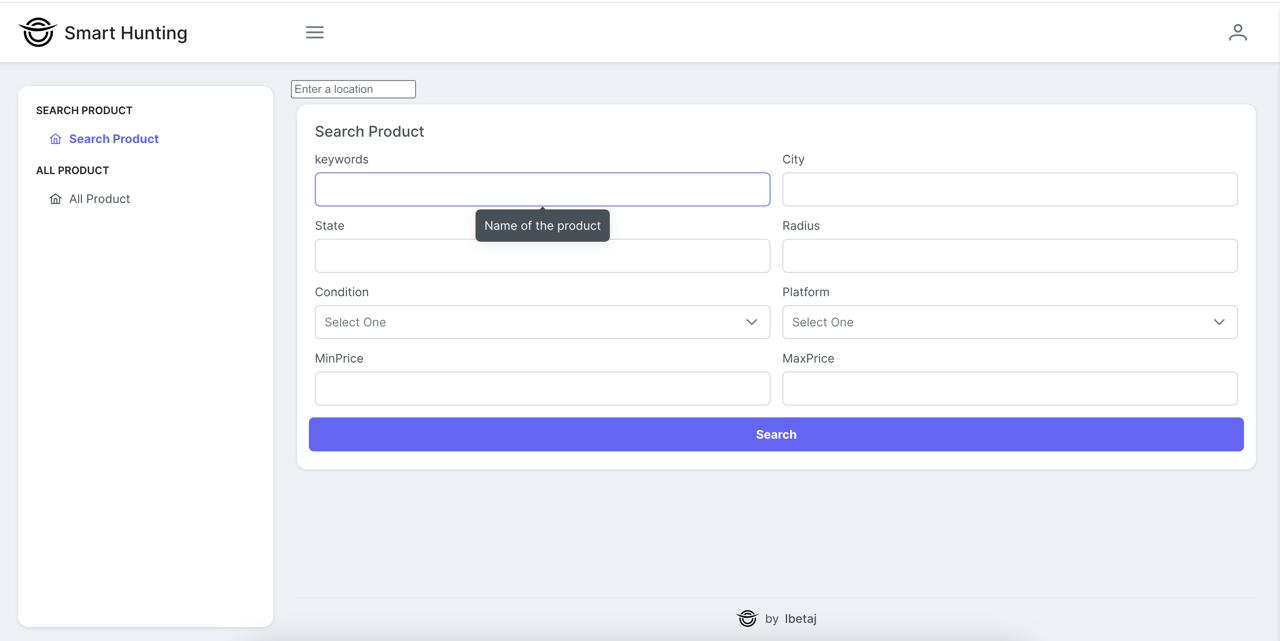


Figure ‎4‑4. Product Search Form

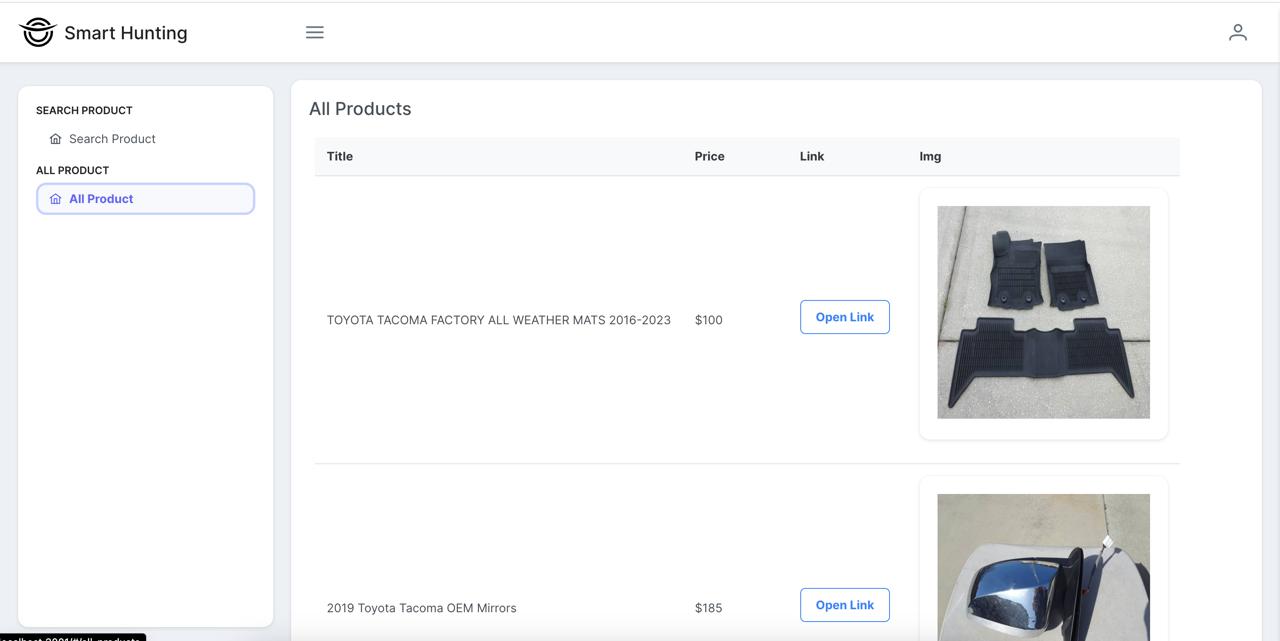


Figure ‎4‑5. PRODUCT LISTING AND MANAGEMENT

## Product Listing and Management:

The application handled scraped data nicely since they were categorized in a product list format. I opined that users could easily look at all the products that they had scraped earlier, which would prove helpful when it comes to the management of their data. This feature also proved highly functional, where the users were able to manipulate product data easily through an interface.

## AI Model Training and Prediction:

The Random Forest Classifier was used to establish the AI model for trend prediction, and for this, 80% of the data was trained and 20% of the data was used for testing. The training process utilized coding in factors, for example, Pandas for handling data and Scikit-Learn for algorithm use. The model gained enough knowledge from the historical information and proved to achieve acceptable accuracy in predicting the status of ‘trending’ or ‘not-trending’ of a product.

## Deployment of the AI Model:

The AI model was deployed with the help of a Flask server, which forms a RESTful API endpoint called “predict,” where product data was being fed from the front end, and then the prediction was made with the help of the trained model, and the results were sent back immediately. The linking of the front-end React components with the Flask server gave users instant feedback, which was well appreciated.

## Prediction Accuracy:

The Random Forest model provided a reasonable result in terms of predicting product trends; the binary output ‘0’ for ‘not trending’ and ‘1’ for ‘trending’ can be useful in real-time decision-making. The accuracy of the model was established from the testing data and clearly showed its ability to analyse product trends and generate useful insights.

In order to enhance the performance of the AI model, there are several ways that can be done, which mainly focus on the data and methods. One of the major concerns is the aspect of data, both in quality and quantity. This shows that proper feature engineering can always be useful in improving the Random Forest’s predictive capability by including other variables that can represent other factors that are not directly observable, such as time features or even other social media metrics. However, a problem of class imbalance where there are more ‘not trending’ products than ‘trending’ products must be handled. Some of the measures that can be used to counter this include oversampling, under sampling, or even using other techniques like SMOTE (Synthetic Minority Oversampling Technique). Lastly, it is also possible to collect more data for the model by either having longer historical data or applying data augmentation techniques to enhance the model’s performance.

It also includes model optimization, which also plays a big part in the overall process. These include items like the number of trees and tree depth, which can greatly affect the model’s performance. Some other possibilities of ensemble methods can be trying to combine Random Forest with Gradient Boosting or XGBoost in a stack model, and it could also lead to increasing the accuracy of the model as well. Factors such as restricted data, which may be sparse or the records may be incomplete, may raise issues with the model generalization. Refinement of data granularity, use of advanced imputation methods to handle missing values, and integration of time-series models can improve the predictive power of the model.

For real-time decision-making, the most important aspect is how the model can be updated with a new data set. It is possible to use online learning algorithms that adapt their model dynamically and therefore prevent the model from becoming outdated. Furthermore, applying methods such as SHAP (Shapley Additive Explanations) to the model will enhance its interpret-ability and thus enable the identification of significant features for consideration in future data gathering and feature extraction. In particular, these strategies, related to the improvement of the data quality and tuning of the Random Forest model, are vital for increasing the predictive accuracy of the product trends.

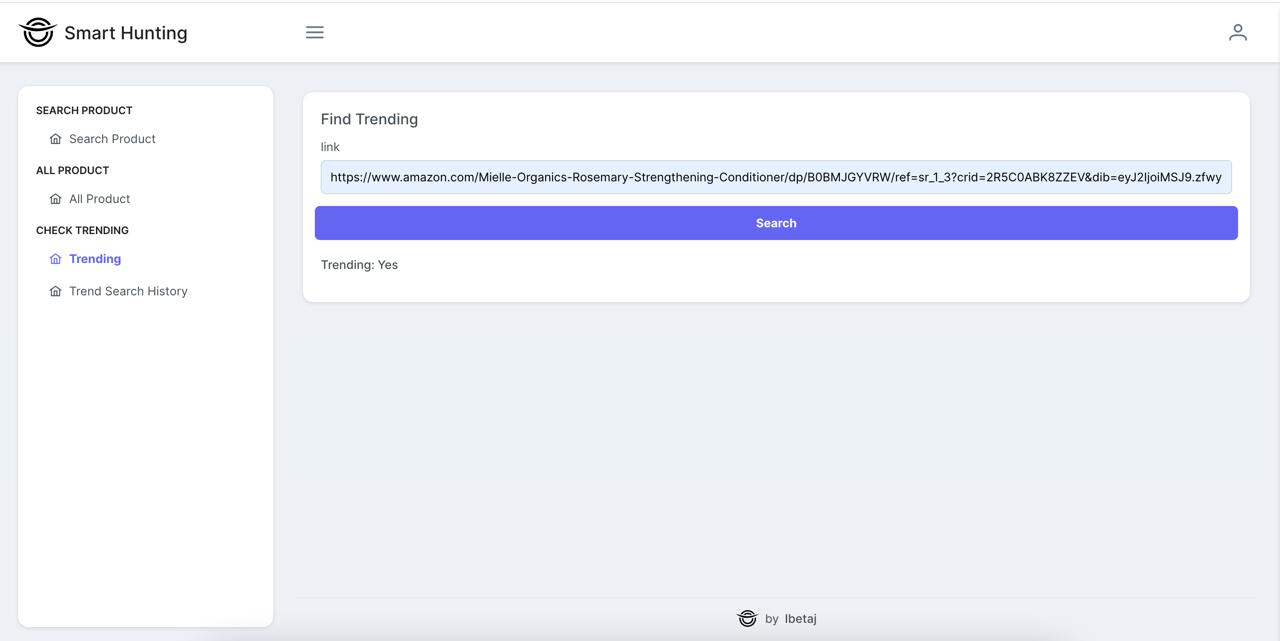


Figure ‎4‑6. Product Search for Trending

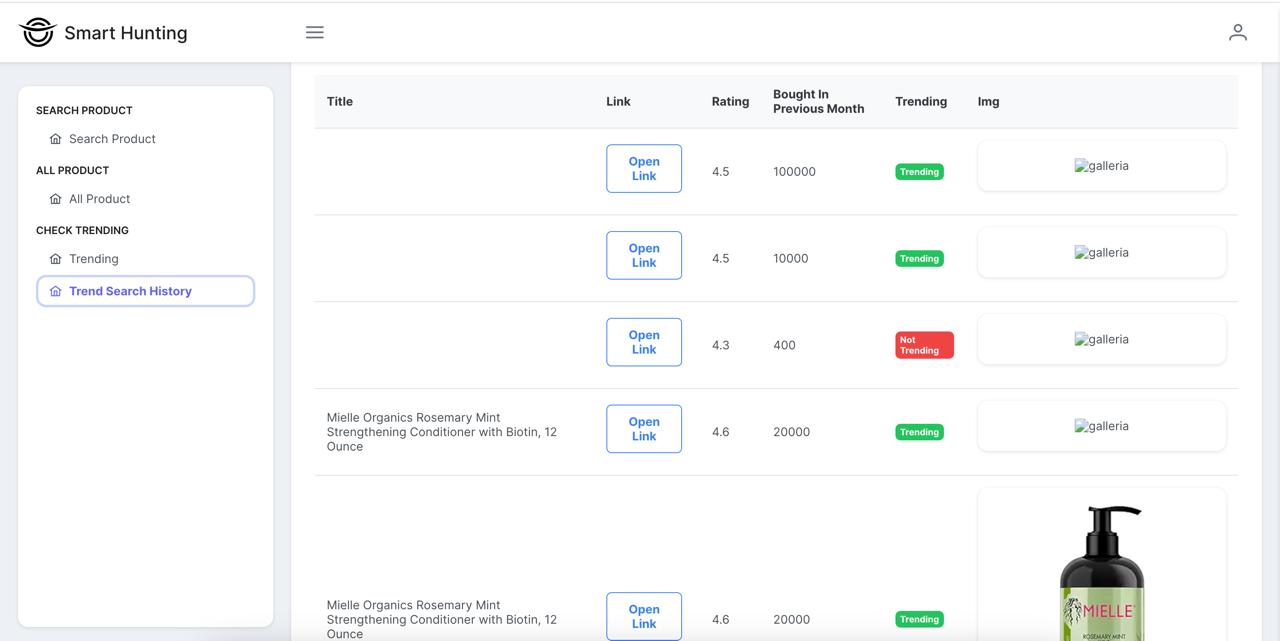


Figure ‎4‑7. Trend Search History

## User Interface Experience:

The application's user interface, developed with React, provided a dynamic and responsive experience. Users were able to easily navigate between different functionalities, such as registration, login, product scraping, and trend analysis. The interface design contributed to a user-friendly experience, enabling quick access to key features and facilitating seamless interaction with the application.

These results collectively confirm the effectiveness of the chosen technologies and methodologies in developing a robust application capable of scraping data, managing product information, and predicting market trends in real-time.

# Chapter 5

# Conclusion:

This study has given a good exploration of smart product hunting on the techniques on the different platforms, especially the Facebook Marketplace and Craigslist. Therefore, the findings underscore the assertion that with the current dynamic e-commerce environment, the adoption of higher computational paradigms such as algorithms, machine learning, and artificial intelligence (AI) is fundamental in improving the product search and recommendation systems. Thus, large-scale market platforming occurs as a continuous process of market expansion and adaptation to user preferences and needs, the solution to which requires innovative approaches to address such challenges as big data, active user engagement, and numerous product offers.

This study is able to show and explain that the change of Facebook, for instance, from being just a site of social networking has shifted market trends and buyers’ behaviour greatly. Currently, the Facebook Marketplace has added new ways of innovation for the sellers and buyers through traded data, including likes, shares, and other searched items. However, user involvement in the platform makes it challenging to recommend the right products, thus the application of strategic techniques with art intelligence for product search.

 Craigslist, on the other hand, resembles a classical classifieds site that has a very simple design that may be somewhat restricting at times but offers millions of services and items. This is where smart product search options on Craigslist become necessary because it is a user-generated content platform, and this contributes to more unstructured platforms than the traditional methods of selling products, which can make this a complicated process since it requires buyers and sellers to build trust. The application of ontologies, natural language processing, and sentiment analysis proves to be valuable in improving the platform’s search functionality and for the end-user.

This study equally sheds light on the importance of technology and the internet in today’s business decision-making process, especially for business entities such as MSMEs. These giant websites contain prospects to avail market opportunities, increase productivity, and direct customer interaction but also include paradoxes like technological hindrances, the existence of digital gaps, and concerns concerning data security and privacy issues. Hence, these tools are equally being used by MSMEs; however, they need to learn how to work with these tools effectively and effectively minimize risks that come with adopting these tools.

Smart product hunting has been established to achieve certain results through the use of AI and big data analytics in enhancing the users’ experience and improving the effectiveness of recommendation systems. Tools like sentiment analysis, natural language processing, and fourth-party ontology-based search optimization offer heightened understanding of users’ behaviour, their preferences, as well as the market status, which in turn facilitates effective marketing techniques.

In addition, they have discussed how this research has revealed the technical process of creating a smart product hunting application based on a React front end and Node back-end. Select Node.js and Python for the back end. The functionalities include registration of users, scraping the products from sites like Facebook and Craigslist, and use of AI in analyzing the trends of products. The discussed case of implementing machine learning algorithms, such as the Random Forest algorithm, for product prediction shows the necessity of the AI’s role in improving the effectiveness and reliability of product recommendations.

This study highlights the need to apply efficient technologies in enhancing cross-sell dynamics for product search and recommendation in digital applications. Therefore, through the use of elaborate equations and statistical methodologies, it becomes easier for enterprises to satisfy the customers’ needs, hence leading to enhanced user satisfaction and enterprise success in the world of e-commerce. Further studies should be thus conducted to propose the improvements of the smart product hunting system for the future needs of e-commerce.

# References:

Abighail, B.M.D. *et al.* (2023) 'Sentiment Analysis E-commerce Review,' *Procedia Computer Science*, 227, pp. 1039–1045. https://doi.org/10.1016/j.procs.2023.10.613.

Aftab, F. *et al.* (2023) 'A Comprehensive Survey on Sentiment Analysis Techniques,' *International Journal of Technology*, 14(6), p. 1288. https://doi.org/10.14716/ijtech.v14i6.6632.

Ahad, M. (2023) 'A REVIEW ARTICLE: USE OF SENTIMENT ANALYSIS IN SOCIAL MEDIA,' *International Journal of Engineering Applied Sciences and Technology*, 7(9), pp. 171–176. https://doi.org/10.33564/ijeast.2023.v07i09.026.

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *EM*, *26*(2), 173–194. https://doi.org/10.1007/s12525-016-0219-0

Amaral, M.A.L. and Djuang, G. (2023) 'Relationship Between Social Influence, Shopping Lifestyle, and Impulsive Buying on Purchase Intention of Preloved Products,' *KINERJA*, 27(1), pp. 91–106. https://doi.org/10.24002/kinerja.v27i1.6635.

Arthur. (2023, May 31). *Data Privacy in E-Commerce*. Retrieved July 23, 2024, from https://heydata.eu/en/magazine/data-privacy-in-e-commerce-challenges-and-best-practices-1

Basani, Y., Sibuea, H. V., Sianipar, S. I. P., & Samosir, J. P. (2019). Application of Sentiment Analysis on Product Review E-Commerce. *Journal of Physics. Conference Series*, *1175*, 012103. https://doi.org/10.1088/1742-6596/1175/1/012103

Bogoslov, I.A. *et al.* (2023) 'GAMIFICATION IN E-COMMERCE: ADVANTAGES, CHALLENGES, AND FUTURE TRENDS,' *Revista Economica*, 75(2), pp. 17–33. https://doi.org/10.56043/reveco-2023-0012.

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). A Topic-Based Bibliometric Review of Computers in Human Behavior: Contributors, Collaborations, and Research Topics. *Sustainability*, *13*(9), 4859. https://doi.org/10.3390/su13094859

D. R. (2018) “Analyzing User Behavior of Social Media”, *International Journal on Future Revolution in Computer Science &amp; Communication Engineering*, 4(2), pp. 11–13. Available at: https://www.ijfrcsce.org/index.php/ijfrcsce/article/view/1160 (Accessed: 23June2024).

D’Annunzio, A. and Russo, A. (2021) 'Intermediaries in the Online Advertising Market,' *SSRN Electronic Journal* [Preprint]. https://doi.org/10.2139/ssrn.3892623.

DATA MINING AND E-COMMERCE: METHODS, APPLICATIONS, AND CHALLENGES. (2008). In *https://core.ac.uk/download/pdf/11784915.pdf*. Retrieved July 23, 2024, from https://core.ac.uk/download/pdf/11784915.pdf

Davison, R.M. *et al.* (2023) 'Pickled eggs: GenerativeAIas research assistant or co‐author?,' *Information Systems Journal*, 33(5), pp. 989–994. https://doi.org/10.1111/isj.12455.

Dornubari, E.A., Blossom, O. and G, O.C. (2023) 'Leveraging on ‘simplicity’ as a principle in the design of interior spaces,' *International Journal of Multidisciplinary Comprehensive Research*, 2(6), pp. 01–06. https://doi.org/10.54660/ijmcr.2023.2.6.01-06.

Dwianto, R.A. *et al.* (2023) 'Facebook Ads, COVID-19, and Smart City Dissemination,' *KnE Social Sciences* [Preprint]. https://doi.org/10.18502/kss.v8i3.12824.

Ettis, S. A., & Abidine, A. B. Z. E. (2017). Consumer Behavior in M-Commerce. In *Advances in e-business research series* (pp. 264–287). https://doi.org/10.4018/978-1-5225-2469-4.ch015

Fan, Z. (2024) 'Research on the Social Media and Online Sales,' *Advances in Economics Management and Political Sciences*, 79(1), pp. 76–81. https://doi.org/10.54254/2754-1169/79/20241886.

Felix, A. and Rembulan, G.D. (2023) 'Analysis of Key Factors for Improved Customer Experience, Engagement, and Loyalty in the E-Commerce Industry in Indonesia,' *Aptisi Transactions on Technopreneurship (ATT)*, 5(2sp), pp. 196–208. https://doi.org/10.34306/att.v5i2sp.350.

Fitri, M., Sirait, F.Y. and Batubara, M. (2024) 'Analisis Dampak Negatif Social Media dan E-Commerce (Social Commerce) Pada Usaha Mikro Kecil Menengah,' *Movere Journal*, 6(1), pp. 70–79. https://doi.org/10.53654/mv.v6i1.421.

Gilbert, S. *et al.* (2023) 'Learning From Experience and Finding the Right Balance in the Governance of Artificial Intelligence and Digital Health Technologies,' *Journal of Medical Internet Research*, 25, p. e43682. https://doi.org/10.2196/43682.

Goddu, J., & Goddu, J. (2024, March 11). *The Importance of Trust in Consumer Behavior*. Sogolytics Blog. https://www.sogolytics.com/blog/importance-of-trust-in-consumer-behavior/

Gurung, A., & Raja, M. (2016). Online privacy and security concerns of consumers. *Information & Computer Security/Information and Computer Security*, *24*(4), 348–371. https://doi.org/10.1108/ics-05-2015-0020

He, Y. *et al.* (2023a) 'Que2Engage: Embedding-based Retrieval for Relevant and Engaging Products at Facebook Marketplace.' https://doi.org/10.1145/3543873.3584633.

He, Y. *et al.* (2023b) 'Que2Engage: Embedding-based Retrieval for Relevant and Engaging Products at Facebook Marketplace.' https://doi.org/10.1145/3543873.3584633.

He, Y. *et al.* (2023c) 'Que2Engage: Embedding-based Retrieval for Relevant and Engaging Products at Facebook Marketplace.' https://doi.org/10.1145/3543873.3584633.

Hou, T. *et al.* (2019) 'Mining customer product reviews for product development: A summarization process,' *Expert Systems With Applications*, 132, pp. 141–150. https://doi.org/10.1016/j.eswa.2019.04.069.

Injadat, M., Salo, F., & Nassif, A. B. (2016). Data mining techniques in social media: A survey. *Neurocomputing*, *214*, 654–670. https://doi.org/10.1016/j.neucom.2016.06.045

Journal, I. (2024) 'Enhancing the Shopping Experience: Leveraging API’s for Intelligent Product Search,' *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 08(02), pp. 1–13. https://doi.org/10.55041/ijsrem28673.

K, Dr.K. (2024) 'Social Media and Social Networking,' *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 08(03), pp. 1–5. https://doi.org/10.55041/ijsrem29635.

Karimova, F. (2016) 'A Survey of e-Commerce Recommender Systems,' *European Scientific Journal*, 12(34), p. 75. https://doi.org/10.19044/esj.2016.v12n34p75.

Kaur, K., Natarajan, S., Kaur, K., Prasanthi, P., Kalaivani, A., Natarajan, S., Kannan, T., & Rajalakshmi, M. (2024). *ARTIFICIAL INTELLIGENCE IN E-COMMERCE: APPLICATIONS, IMPLICATIONS, AND CHALLENGES*.

Kennedy, M. (2023, May 3). Trust and Safety in Online Marketplaces. *KOALA*. https://www.go-koala.com/timeshareblog/trust-and-safety-in-online-marketplaces/

Khan, M. H. (2023, July 20). *The Impact of Classified Ads Websites on Local Economies  - Blog*. Blog. https://directorist.com/blog/classified-ads-websites-on-local-economies/

Komala, M.G. *et al.* (2023) 'Investigating the Regulatory Process, Safety, Efficacy and Product Transparency for Nutraceuticals in the USA, Europe and Australia,' *Foods*, 12(2), p. 427. https://doi.org/10.3390/foods12020427.

Lafia, S., Million, A.J. and Hemphill, L. (2023) 'Direct, Orienting, and Scenic Paths: How Users Navigate Search in a Research Data Archive.' https://doi.org/10.1145/3576840.3578275.

Liu, B., & Zhang, L. (2012). A Survey of Opinion Mining and Sentiment Analysis. In *Springer eBooks* (pp. 415–463). https://doi.org/10.1007/978-1-4614-3223-4\_13

Liu, J. *et al.* (2023) 'JDsearch: A Personalized Product Search Dataset with Real Queries and Full Interactions.' https://doi.org/10.1145/3539618.3591900.

Luca, M. (2017). Designing Online Marketplaces: Trust and Reputation Mechanisms. *Innovation Policy and the Economy*, *17*, 77–93. https://doi.org/10.1086/688845

M, H. K. S., Hegde, S., G, S., M, S., R, S., & N, S. L. (2021). *User Interest Prediction based on Social Network Profile with Machine Learning*. https://doi.org/10.1109/i2ct51068.2021.9418126

Mccauley, J., Shenker, S. and Varghese, G. (2023) 'Extracting the Essential Simplicity of the Internet,' *Communications of the ACM*, 66(2), pp. 64–74. https://doi.org/10.1145/3547137.

Miranda, V. E. (2022). Exploring Financial Risk Management for E-Commerce Startups. *American International Journal of Business Management (AIJBM)*, *5*(12), 112–123.

Napitupulu, R., Asti, W. and Hermawan, P. (2023) 'Developing a Business Model Canvas for Online Classified Ads Platform: A Startup Case Study,' *International Journal of Current Science Research and Review*, 06(06). https://doi.org/10.47191/ijcsrr/v6-i6-31.

Niu, S. *et al.* (2023) 'Building Credibility, Trust, and Safety on Video-Sharing Platforms.' https://doi.org/10.1145/3544549.3573809.

Ogundipe, N.D.O., Babatunde, N.S.O. and Abaku, N.E.A. (2024) 'AI AND PRODUCT MANAGEMENT: A THEORETICAL OVERVIEW FROM IDEA TO MARKET,' *International Journal of Management & Entrepreneurship Research*, 6(3), pp. 950–969. https://doi.org/10.51594/ijmer.v6i3.965.

Okhrem, P. (2024, May 17). *Dynamic Pricing in Ecommerce: How It Works*. Elogic. https://elogic.co/blog/dynamic-pricing-in-ecommerce/

Oktaviani, N.A., Maulana, N.A. and Firmansyah, N.R. (2023) 'Peranan Media Sosial Facebook dalam Meningkatkan Komunikasi Pemasaran di Era Digital,' *MUKASI Jurnal Ilmu Komunikasi*, 2(2), pp. 143–150. https://doi.org/10.54259/mukasi.v2i2.1592.

Pal, V.B. and Kumari, P. (2023) 'Consumer Buying Behaviour towards Online Shopping,' *Journal of Business Management and Information Systems*, 10(1), pp. 4–9. https://doi.org/10.48001/jbmis.2023.1001002.

Park, J.S. *et al.* (2023) 'Generative Agents: Interactive Simulacra of Human Behavior.' https://doi.org/10.1145/3586183.3606763.

Park, Y., McCoy, D., & Shi, E. (2017). Understanding Craigslist Rental Scams. In *Lecture notes in computer science* (pp. 3–21). https://doi.org/10.1007/978-3-662-54970-4\_1

Phan, C.T.T., Đặng, Q.T. and Nguyễn, T.L. (2023) 'Belief of Customers in Social Commerce Performed via Social Networking Sites: An Empirical Study from Hồ Chí Minh City,' *Vietnam Academy of Social Sciences S VSSR*, 214(2), pp. 61–80. https://doi.org/10.56794/vssr.2(214).61-80.

Quan, H. *et al.* (2023a) 'Big Data and AI-Driven Product Design: A Survey,' *Applied Sciences*, 13(16), p. 9433. https://doi.org/10.3390/app13169433.

Quan, H. *et al.* (2023b) 'Big Data and AI-Driven Product Design: A Survey,' *Applied Sciences*, 13(16), p. 9433. https://doi.org/10.3390/app13169433.

Ramos, A.A., Arcilla, I.J.A. and Macatangay, C.J.C. (2023) 'DOWNLOAD PDF DOWNLOAD XML FACTORS AFFECTING CUSTOMER’S SATISFACTION ON FACEBOOK MARKETPLACE AS PURCHASING PLATFORM: EVIDENCE FROM LAGUNA, PHILIPPINES,' *Deleted Journal*, 1(1), pp. 26–29. https://doi.org/10.26480/jhcdc.01.2023.26.29.

Raorane, A., & Kulkarni, R. V. (2011). Data Mining Techniques: A Source for Consumer Behavior Analysis. *International Journal of Database Management Systems*, *3*(3), 45–56. https://doi.org/10.5121/ijdms.2011.3304

S, B.N. *et al.* (2023) 'Opinion Mining For E-Commerce in Social Media Accounts,' *International Journal for Research in Applied Science and Engineering Technology*, 11(3), pp. 819–826. https://doi.org/10.22214/ijraset.2023.49535.

Saikia, W. and Bhattacharjee, A. (2023) 'Digital consumer engagement in a social network: A literature review applying TCCM framework,' *International Journal of Consumer Studies*, 48(1). https://doi.org/10.1111/ijcs.12981.

Saputri, S.A. *et al.* (2023) 'PERAN MARKETPLACE DALAM MENINGKATKAN DAYA SAING UMKM DI INDONESIA,' *KNOWLEDGE Jurnal Inovasi Hasil Penelitian Dan Pengembangan*, 3(1), pp. 69–75. https://doi.org/10.51878/knowledge.v3i1.2199.

Setiadi, B. *et al.* (2023) 'The Application of Delone and Mclean Framework to Analyze the Relationship Between Customer Satisfaction and User Experience of Mobile Application,' *Jurnal Sistim Informasi Dan Teknologi*, pp. 84–89. https://doi.org/10.37034/jsisfotek.v5i1.207.

Siahaan, A.M., Siboro, D.T. and Saragi, D.R.R. (2024) 'Strategi Meningkatkan Profitabilitas Dengan Memakai Penjualan Berbasis Digital Pada Usaha Mikro Kecil Menengah (UMKM) Di Kota Medan,' *Jurnal Akuntansi Bisnis*, 22(1), pp. 29–40. https://doi.org/10.24167/jab.v22i1.11112.

Simoben, C.V. *et al.* (2023) 'Challenges in natural product-based drug discovery assisted with in silico-based methods,' *RSC Advances*, 13(45), pp. 31578–31594. https://doi.org/10.1039/d3ra06831e.

Singh, M. and Singh, G. (2018) 'Impact of social media on e-commerce,' *International Journal of Engineering & Technology*, 7(2.30), p. 21. https://doi.org/10.14419/ijet.v7i2.30.13457.

Sisodiya, R. (2024) 'E-Commerce Recommendation System,' *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 08(04), pp. 1–5. https://doi.org/10.55041/ijsrem30996.

Srivastava, S. and Nalla, B. (2023) 'A Review –of Sales Encourage Online Consumer selected FMCG Product,' *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 07(11), pp. 1–11. https://doi.org/10.55041/ijsrem27238.

Stige, Å. *et al.* (2023) 'Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda,' *Information Technology and People* [Preprint]. https://doi.org/10.1108/itp-07-2022-0519.

Sudirman, L. and Disemadi, H.S. (2023) 'THE ROLE OF INDONESIAN ONLINE MARKETPLACES IN INTELLECTUAL PROPERTY RIGHTS INFRINGEMENTS: A COMPARATIVE ANALYSIS,' *Jurnal Pembaharuan Hukum*, 10(1), p. 90. https://doi.org/10.26532/jph.v10i1.30883.

Sugiharto, B.H. (2024) 'The Role of E-commerce for MSMEs as a Digital Marketing Strategy in Facing Industrial Revolution 4.0,' *Deleted Journal*, 1(1), pp. 99–107. https://doi.org/10.62207/80ndq458.

Taherdoost, H., & Madanchian, M. (2023). Blockchain-Based E-Commerce: A Review on Applications and Challenges. *Electronics*, *12*(8), 1889. https://doi.org/10.3390/electronics12081889

Tarmizi, R. *et al.* (2023) 'Harnessing Digital Platforms for Entrepreneurial Success: A Study of Technopreneurship Trends and Practices,' *Aptisi Transactions on Technopreneurship (ATT)*, 5(3), pp. 278–290. https://doi.org/10.34306/att.v5i3.360.

Van Etten, J. *et al.* (2023) 'Data-driven approaches can harness crop diversity to address heterogeneous needs for breeding products,' *Proceedings of the National Academy of Sciences*, 120(14). https://doi.org/10.1073/pnas.2205771120.

Vavliakis, K. N., Katsikopoulos, G., & Symeonidis, A. L. (2019). E-commerce Personalization with Elasticsearch. *Procedia Computer Science*, *151*, 1128–1133. https://doi.org/10.1016/j.procs.2019.04.160

Viswanath, B., Mislove, A., Cha, M., & Gummadi, K. P. (2009). *On the evolution of user interaction in Facebook*. https://doi.org/10.1145/1592665.1592675

Warner, R. and Sherbia, P.J. (2023) 'A Niche Discovery in the Metaverse,' *International Journal of Business & Management Studies*, 04(02), pp. 08–19. https://doi.org/10.56734/ijbms.v4n2a2.

Xu, L. and Sang, X. (2022) 'E-Commerce Online Shopping Platform Recommendation Model Based on Integrated Personalized Recommendation,' *Scientific Programming*, 2022, pp. 1–9. https://doi.org/10.1155/2022/4823828.

Xu, Q.A., Chang, V. and Jayne, C. (2022) 'A systematic review of social media-based sentiment analysis: Emerging trends and challenges,' *Decision Analytics Journal*, 3, p. 100073. https://doi.org/10.1016/j.dajour.2022.100073.

Yadav, J. (2023) 'Sentiment Analysis on Social Media.' https://doi.org/10.32388/yf9x04.

Ypanaqué, B.L.M. and Dueñas, A.M.H. (2023) 'Review of Sentiment Analysis Techniques in Social Networks,' *Revista De Investigación De Sistemas E Informática*, 16(2), pp. 189–201. https://doi.org/10.15381/risi.v16i2.27144.

Zafarani, R., Abbasi, M.A. and Liu, H. (2014) *Social Media Mining*. Cambridge University Press.

Zhang, Q., Lu, J., & Jin, Y. (2020). Artificial intelligence in recommender systems. *Complex & Intelligent Systems*, *07*, 439–457. https://doi.org/10.1007/s40747-020-00212-w