

VIETNAM NATIONAL UNIVERSITY, HO CHI MINH CITY
UNIVERSITY OF INFORMATION TECHNOLOGY
FACULTY OF INFORMATION SYSTEMS



FINAL PROJECT REPORT
ENTERPRISE RESOURCE PLANNING

TOPIC:

**APPLY SEMANTIC SEARCH AND RECOMMENDER SYSTEM
IN E-COMMERCE WEBSITE USING CLIP MODEL**

Instructors:

Dr. Tran Van Hai Trieu

Engineer. Le Vo Dinh Kha

Group 1:

1	Vo Thi Thu Tien	21520482
2	Nguyen Lam Nhat Anh	21521832
3	Dang Thanh Dat	21521928
4	Ngo Thuan Phat	21522445

HO CHI MINH CITY, MAY 2024

INSTRUCTOR'S FEEDBACK

Ho Chi Minh City, June 12, 2024

ASSIGNMENT AND MEMBER EVALUATION TABLE

Name	Student ID	Assignment	Evaluation
Dang Thanh Dat	21521928	<p>Week 1: Collect and review academic papers and case studies related to the topic.</p> <p>Week 2: Analyze the method for processing text data, how to convert text to vector.</p> <p>Week 3: Research semantic search engine in existing e-commerce platforms and develop the model for our searching project.</p> <p>Week 4: Improve the semantic analysis models for relevance information.</p> <p>Week 5: Test the system and write the final report.</p>	<p>Week1: 9/10</p> <p>Week 2: 10/10</p> <p>Week 3: 8/10</p> <p>Week 4: 8/10</p> <p>Week 5: 9/10</p>
Nguyen Lam Nhat Anh	21521832	<p>Week 1: Collect and review academic papers and case studies related to the topic.</p> <p>Week 2:</p>	<p>Week1: 10/10</p> <p>Week 2: 9/10</p>

		<p>Clean and preprocess the collected data of product information.</p> <p>Week 3:</p> <p>Study vector database platform for storing our database and meta data</p> <p>Week 4:</p> <p>Set up data storage and retrieval infrastructure for deployment.</p> <p>Week 5:</p> <p>Deploy the system in AWS to get an API for our E-Commerce platform.</p>	<p>Week 3:</p> <p>8/10</p> <p>Week 4:</p> <p>8/10</p> <p>Week 5:</p> <p>10/10</p>
Vo Thi Thu Tien	21520482	<p>Week 1:</p> <p>Investigate customer needs and preferences in the context of product recommendations.</p> <p>Week 2:</p> <p>Research semantic analysis algorithms to understand product descriptions and reviews.</p> <p>Week 3:</p> <p>Research how to convert image data to vector and store in database.</p> <p>Week 4:</p>	<p>Week1:</p> <p>9/10</p> <p>Week 2:</p> <p>10/10</p> <p>Week 3:</p> <p>8/10</p> <p>Week 4:</p> <p>8/10</p>

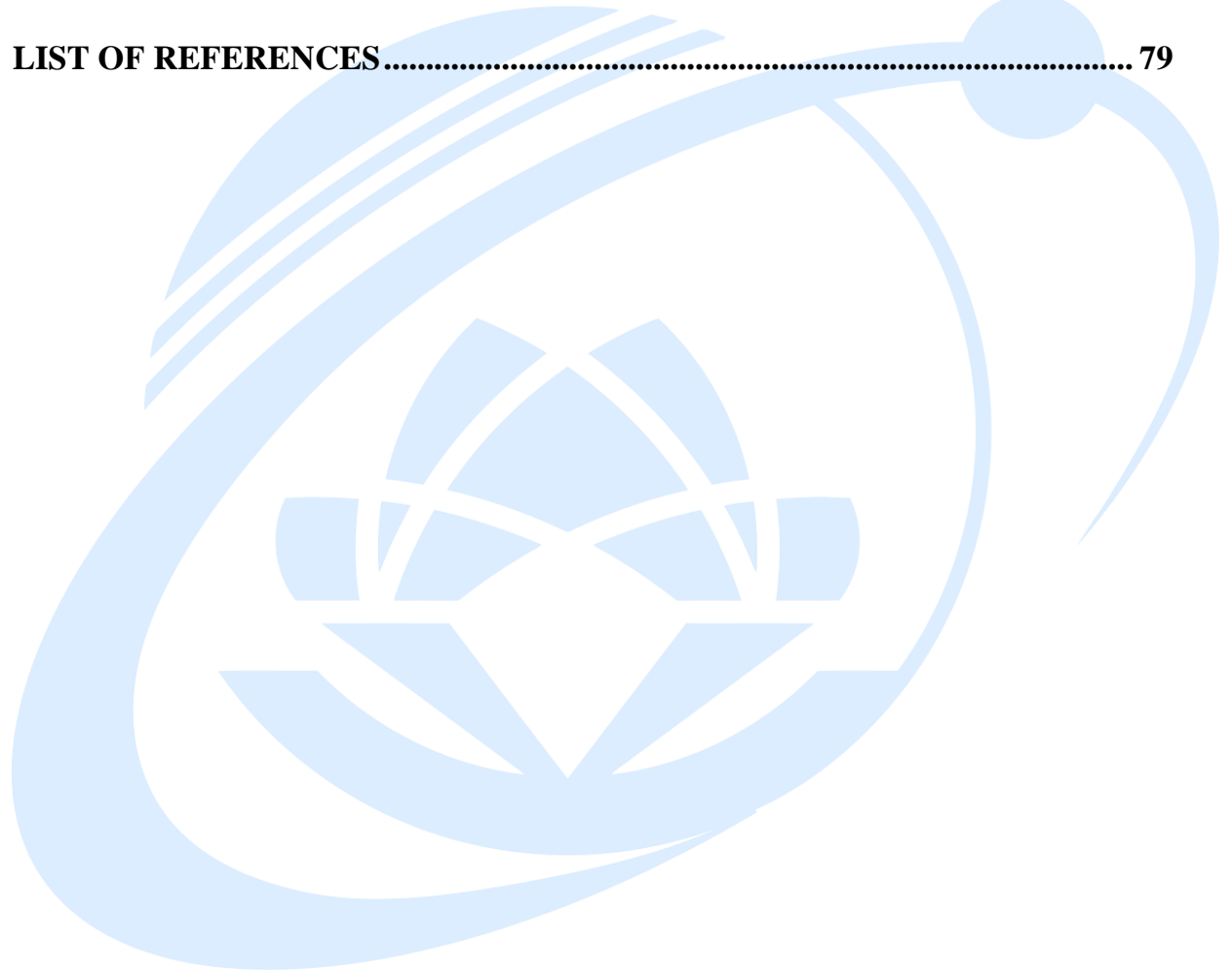
		<p>Integrate optimized models into a cohesive system and test for consistency.</p> <p>Week 5:</p> <p>Develop a comprehensive report on research findings and system performance.</p>	<p>Week 5:</p> <p>9/10</p>
<p>Ngo Thuan Phat</p>	<p>21522445</p>	<p>Week 1:</p> <p>Investigate customer needs and preferences in the context of product recommendations.</p> <p>Week 2:</p> <p>Perform image preprocessing tasks like resizing and normalization.</p> <p>Week 3:</p> <p>Research how to convert image data to vector and store in database.</p> <p>Week 4:</p> <p>Analyze how to store vector including image vector and text vector in vector database.</p> <p>Week 5:</p> <p>Perform end-to-end testing of the entire system.</p>	<p>Week1:</p> <p>8/10</p> <p>Week 2:</p> <p>10/10</p> <p>Week 3:</p> <p>8/10</p> <p>Week 4:</p> <p>9/10</p> <p>Week 5:</p> <p>9/10</p>

TABLE OF CONTENTS

LIST OF FIGURES	9
LIST OF TABLES	10
LIST OF ABBREVIATIONS	11
PROJECT SUMMARY	12
ACKNOWLEDGEMENTS	13
CHAPTER 1. GENERAL INTRODUCTION	14
1.1. ERP project overview.....	14
1.2. Project scope.....	16
1.3. Project objectives	17
1.4. Stakeholders and key personnel	18
1.5. Report layout	21
1.6. Project plan – WBS	22
1.7. Work assignment table	30
CHAPTER 2: THEORETICAL OVERVIEW	36
2.1. ERP introduction	36
2.2. Technologies used in system building	37
2.1.1. Overview of technology.....	37
2.1.2. History of development.....	37
2.3. Description of the problem applying ERP to the project topic (or Case study).....	39
2.3.1. Hybrid search using image vision and semantic of text	39
2.3.2. Case study	40
CHAPTER 3: BUSINESS PROCESS SYSTEM	42
3.1. System overview	42
3.2. Procurement process.....	43
3.2.1. BPMN	43
3.2.2. Specification.....	43
3.3. Sales process.....	46

3.3.1. BPMN	46
3.3.2. Specification.....	46
3.4. Inventory receiving process.....	47
3.4.1. BPMN	47
3.4.2. Specification.....	47
3.5. Inventory shipping process.....	48
3.5.1. BPMN	48
3.5.2. Specification.....	49
3.6. Financial and Accounting management process	50
3.6.1. BPMN	50
3.6.2. Specification.....	51
3.7. Human resources management process.....	52
3.7.1. BPMN	52
3.7.2. Specification.....	53
3.8. Customer service process	55
3.8.1. BPMN	55
3.8.2. Specification.....	56
3.9. AI implementation.....	58
3.9.1. BPMN	58
3.9.2. Specification.....	59
CHAPTER 4: PROPOSED ALGORITHMS AND SOLUTIONS.....	61
4.1. Introduction to dataset.....	61
4.1.1 Data collection	61
4.2. Introduces the application of deep learning to the process	64
4.3. Implementation steps.....	65
4.3.1. Combine retrieval using clip-ViT-B-32 model and BM25 Retriever.....	66

4.3.2. Process the output vector and store in vector database.....	68
4.3.3. Processing data after retrieval	69
4.3.4. Optimize search engine.....	70
CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS.....	75
5.1. Advantages	75
5.2. Limitations.....	76
5.3. Future directions.....	76
LIST OF REFERENCES.....	79



LIST OF FIGURES

Figure 1: WBS	23
Figure 2: WBS - Initiation and Planning phases	24
Figure 3: WBS - Execution phases	25
Figure 4: WBS - Monitoring and Closeout phases	29
Figure 5: BPMN - Procurement process.....	43
Figure 6: BPMN - Sale process	46
Figure 7: BPMN - Inventory receiving process.....	47
Figure 8: Inventory shipping process	49
Figure 9: BPMN - Financial and Accounting management process	50
Figure 10: BPMN - Recruitment process	52
Figure 11: Human resources management process	53
Figure 12: BPMN - Customer service process	56
Figure 13: BPMN - AI implementation.....	58
Figure 14: Store vector database process.....	59
Figure 15: Data set.....	62
Figure 16: Delete data line that has missing values.....	63
Figure 17: Convert all text to lower characters	64
Figure 18: Delete the special characters	64
Figure 19: Architecture of AI model and how the user interact with customer	66
Figure 20: Convert from raw data to vectors.....	67
Figure 21: Fit BM25 model with our data.....	67
Figure 22: Sparse vector of each data point.....	67
Figure 23: CLIP-ViT-B-32 - convert the image to dense vector.....	68
Figure 24:Store data in vector database Pinecone	68
Figure 25: Data that stored in Pinecone including vector and metadata	69
Figure 26: Architecture that customer interact with system.....	69

LIST OF TABLES

Table 1: Stakeholders and key personnel	18
Table 2: Work assignment	30
Table 3: Procurement process specification	43
Table 4: Sale process specification.....	46
Table 5: Inventory receiving process specification	47
Table 6: Inventory shipping process specification	49
Table 7: Human resources management process specification	51
Table 8: Recruitment process	53
Table 9: Human resources management process specification	55
Table 10: Customer service process specification.....	56
Table 11: AI implementation specification	59

LIST OF ABBREVIATIONS

NO.	ABBREVIATION	CONTENT
1	ERP	Enterprise Resource Planning
2	CRM	Customer Relationship Management
3	WBS	Work Breakdown Structure
4	IBDSS	International Brand Distributor Store System
5	UI	User Interface
6	UX	Use Experience
7	AI	Artificial Intelligence
8	IoT	Internet of Things
9	CLIP	Contrastive Language-Image Pre-Training
10	POS	Point of sale
11	CRM	customer relationship management
12	HR	Human Resources
13	PCI	payment card industry
14	BPMN	Process Model and Notation
15	SQL	Structured Query Language
16	RFP	Request for proposal
17	RFQ	Request for quotation
18	HTTP	Hypertext Transfer Protocol
19	HTML	HyperText Markup Language
20	NLP	Natural Language Processing

PROJECT SUMMARY

The International Brand Distributor Store System is a cutting-edge e-commerce platform designed to revolutionize the online retail landscape for clothing. This project addresses the growing need for providing the sale of international fashion brands but also enhances the overall shopping experience through advanced technological features.

In today's competitive market, customers expect more than just a transactional relationship with online retailers. They seek personalized shopping experiences that cater to their unique preferences and habits. The International Brand Distributor Store System aims to bridge this gap by incorporating a recommendation function that leverages customer data to deliver product suggestions.

This recommendation function utilizes advanced algorithms to analyze customer behavior, including browsing history, purchase patterns, and other relevant interactions. By understanding individual customer preferences, the system can suggest products that are more likely to resonate with each user, thereby enhancing customer satisfaction and increasing sales potential.

Moreover, the platform is built with scalability in mind. As the business grows and the product catalog expands, the system can handle increased traffic and a larger volume of transactions without compromising performance. This ensures that the platform remains robust and reliable, even during peak shopping seasons.

In conclusion, the International Brand Distributor Store System represents a significant advancement in the realm of online retail. By focusing on personalized recommendations, scalability, and security, the platform is poised to set a new standard for e-commerce solutions. This project not only addresses the current needs of online retailers but also anticipates future trends, ensuring long-term success and customer satisfaction.

ACKNOWLEDGEMENTS

We would like to express our deepest gratitude to all those who contributed to the successful completion of our project, the International Brand Distributor Store System.

First and foremost, we extend our heartfelt thanks to our instructors, **Dr. Tran Van Hai Trieu** and **TA. Le Vo Dinh Kha**, for their invaluable guidance, support, and encouragement throughout this project. Their expertise and insightful feedback were instrumental in shaping our work and ensuring its quality. We appreciate their patience and dedication in helping us navigate the complexities of this project.

We are also grateful to our peers for their constant support and constructive criticism, which motivated us to improve and refine our system. Their collaborative spirit and shared knowledge significantly enriched our learning experience.

Furthermore, we acknowledge the support of our families and friends, whose encouragement and understanding provided us with the strength and determination to persevere through the challenges we encountered.

Lastly, we would like to thank all the individuals and organizations who participated in our surveys and interviews. Their willingness to share their experiences and insights was crucial in shaping the requirements and functionalities of our system.

Thank you all for your unwavering support and contribution to this project.

Sincerely,

Group 1

CHAPTER 1. GENERAL INTRODUCTION

In the first chapter, a brief overview of the topic is presented. From there, objectives are set and solutions for the problem are proposed, as well as identifying the target audience and stakeholders involved with the topic.

1.1. ERP project overview

Nowadays, Enterprise Resource Planning (ERP) systems are integrated software platforms that manage and automate a wide range of business processes within an organization. ERP systems centralize data and processes from various departments, such as finance, human resources, sales, procurement, and inventory management, into a unified system. This integration provides real-time visibility, improves efficiency, reduces errors, and supports data-driven decision-making. ERP systems are designed to enhance productivity, streamline operations, and facilitate better communication and collaboration across an organization.

Application of ERP in a Clothing Store

General Clothing Store:

Inventory Management: ERP systems help clothing stores maintain optimal inventory levels by tracking stock in real-time, managing product variants (sizes, colors, styles), and automating reorder processes. This ensures that popular items are always in stock and reduces overstock of slow-moving items.

Sales Management: Integrated POS systems within the ERP allow seamless tracking of sales across all channels (in-store, online, mobile). The ERP system also supports customer relationship management (CRM) by tracking customer preferences and purchase history, enabling personalized marketing and improved customer service.

Procurement: Automating the procurement process, ERP systems streamline supplier management, purchase orders, and inventory replenishment. This ensures timely restocking and optimal supplier performance.

Financial Management: ERP systems manage all financial transactions, including accounts payable and receivable, expense tracking, and financial reporting. This ensures accurate financial records and compliance with regulations.

Human Resources: ERP systems streamline HR processes such as payroll, recruitment, performance management, and training. This enhances employee management and productivity.

Customer Service: ERP systems provide tools for efficient customer support, including ticketing systems, knowledge bases, and feedback collection. This improves customer satisfaction and loyalty.

International Brand Distributor Store System:

Inventory Management: For an international brand distributor, ERP systems handle complex inventory management needs, including tracking inventory across multiple locations and ensuring compliance with international shipping regulations. Real-time inventory updates help prevent stockouts and overstock situations.

Sales Management: ERP systems support global sales operations by integrating multiple sales channels and managing international transactions. The CRM component helps track customer interactions and preferences globally, enabling targeted marketing campaigns and personalized recommendations.

Procurement: The ERP system streamlines global procurement processes, managing international suppliers, purchase orders, and compliance with customs regulations. Automated procurement processes ensure timely and cost-effective restocking of inventory.

Financial Management: Managing finances on a global scale, the ERP system handles multi-currency transactions, international tax regulations, and comprehensive financial reporting. This ensures financial accuracy and regulatory compliance across all regions.

Human Resources: For international operations, the ERP system manages a diverse workforce across different regions, streamlining payroll, recruitment, and performance management processes while ensuring compliance with local labor laws.

Customer Service: ERP systems enhance global customer service capabilities by providing tools for efficient issue resolution, customer feedback analysis, and multi-

language support. This ensures a consistent and high-quality customer experience worldwide.

ERP systems offer a solution to manage and optimize various business processes, providing significant benefits to both general clothing stores and international brand distributor store systems. By integrating key functions such as inventory management, sales, procurement, financial management, human resources, and customer service, ERP systems improve efficiency, enhance customer satisfaction, and support business growth. Whether for a local clothing store or a global brand distributor, ERP systems play a crucial role in achieving operational excellence and competitive advantage.

1.2. Project scope

The scope of the project is based on the YOLOM store (international brand distributor store system) needs and expectations or our customer's requirements. It is presented as listed statement below:

1. Website design and development:

- + Create a responsive and visually appealing website design that aligns with the brand identity of the international clothing brands being distributed.
- + Develop a user-friendly and intuitive website interface that facilitates easy navigation and product discovery.
- + Implement a secure and reliable e-commerce platform to enable secure online transactions and order processing.

2. Product catalog management:

- + Develop a product catalog system to manage product information, including descriptions, images, pricing, and availability.
- + Integrate product data feeds from international clothing brands to ensure accurate and up-to-date product information.
- + Implement effective product search and filtering functionalities to help customers quickly find the products they are looking for.

3. Order management:

- + Develop a robust order management system to track orders, process payments, and manage shipping and delivery.
- + Implement automated email notifications to keep customers informed about order status and delivery updates.

4. Payment management:

- + Integrate secure payment gateways to accept a variety of payment methods, including credit cards, debit cards, and digital wallets.
- + Implement secure payment processing protocols to protect customer financial information.
- + Ensure compliance with relevant payment card industry (PCI) security standards.

1.3. Project objectives

The primary objective of the International Brand Distributor Store System project is to develop a user-friendly e-commerce platform that facilitates the distribution and sale of international brands to Vietnamese customers. This system aims to streamline operations, enhance customer experience, and drive sales growth by leveraging advanced technology and innovative features. One of the standout features of this platform is the recommendation function, which is designed to personalize the shopping experience by suggesting products that customers may be interested in, based on their browsing history, purchase patterns, and preferences.

To achieve this objective, the project will focus on several key areas. Firstly, the platform will integrate a recommendation engine powered by machine learning algorithms. This engine will analyze vast amounts of customer data in real-time to generate personalized product recommendations. By understanding customer behavior and preferences, the system will not only enhance the shopping experience but also increase the likelihood of purchases, thereby boosting overall sales.

The recommendation function will be prominently featured throughout the platform, from the homepage to product pages and checkout processes. When customers log in, they

will be greeted with a curated selection of products tailored to their interests. As they browse, the system will dynamically update recommendations based on their interactions, ensuring that the suggestions remain relevant and enticing. This feature will also include cross-selling and up-selling strategies, presenting complementary products or premium options that align with the customer's preferences and past behaviors.

Moreover, the recommendation engine will continuously learn and improve from customer interactions, becoming more accurate and effective over time. This continuous improvement will be facilitated by advanced data analytics and feedback mechanisms, which will refine the algorithms to better predict customer needs and interests. By delivering highly relevant product suggestions, the system aims to increase customer satisfaction and loyalty, encouraging repeat visits and purchases.

In addition to the recommendation function, the platform will support a wide range of international brands, offering customers a diverse selection of products across various categories. The system will include product descriptions, high-quality images, and customer reviews to aid in decision-making.

In summary, the International Brand Distributor Store System project aims to revolutionize the e-commerce landscape by offering a personalized, efficient, and secure shopping experience. The recommendation function, a cornerstone of the platform, will play a pivotal role in engaging customers and driving sales by presenting them with products that match their interests and preferences. Through advanced technology and a customer-centric approach, the platform will position itself as a leading destination for international brand shopping, fostering growth and success in the global market.

1.4. Stakeholders and key personnel

Table 1: Stakeholders and key personnel

NO	NAME	ROLES AND RESPONSIBILITIES	CONTACT
1	Vo Thi Thu Tien	Role: Scrum Master Responsibilities: + Acilitating Scrum ceremonies (daily stand-ups, sprint planning,	Email: 21520482@gm.uit.edu.vn Phone: 0378722011

		<p>sprint reviews, retrospectives) and ensuring they are productive and time efficient.</p> <ul style="list-style-type: none"> + Identifying and removing any obstacles that may hinder the Development Team's progress. + Educating and coaching the team and organization on Scrum principles and practices to improve efficiency and effectiveness. 	
2	Nguyen Lam Nhat Anh	<p>Role: Product Owner</p> <p>Responsibilities:</p> <ul style="list-style-type: none"> + Creating, prioritizing, and maintaining the product backlog to ensure it reflects the product goals and stakeholder needs. + Defining and communicating the product vision and roadmap to the team and stakeholders. + Acting as the primary point of contact between stakeholders (customers, business leaders, etc.) and the Development Team, gathering requirements and feedback. + Defining the acceptance criteria for user stories and ensuring the product meets these criteria before release. 	<p>Email: 21521832@gm.uit.edu.vn</p> <p>Phone: 0378721998</p>

3	Dang Thanh Dat	Role: AI Engineer Responsibilities: <ul style="list-style-type: none"> + Conducting research into new AI algorithms and techniques. + Developing AI models and applications. Exploring the potential of AI to solve real-world problems. <ul style="list-style-type: none"> + Monitoring the performance of AI systems. + Troubleshooting problems with AI systems. + Updating and improving AI systems over time. 	Email: 21521928@gm.uit.edu.vn Phone: 0926542213
4	Ngo Thuan Phat	Role: Project Manager Responsibilities: <ul style="list-style-type: none"> + Defining the project scope, objectives, and timeline. + Allocating resources, managing budgets, and ensuring the team has what they need to succeed. + Identifying, assessing, and mitigating project risks. + Keeping stakeholders informed about project progress, changes, and any issues that arise. + Tracking project progress and performance against the plan and reporting on status. 	Email: 21522445@gm.uit.edu.vn Phone: 036532995
5	YOLOM Store	Role: Customer Responsibilities:	Email: yolom@gmail.com Phone: 0258888999

		+ Providing the requirements and expectations Product Owner + Participate in Sprint Review meetings to see the work completed during the sprint, provide feedback and discuss any changes or additional requirements + Validate that the delivered work meets the acceptance criteria and business requirements, providing formal acceptance or raising issues that need to be addressed.	
--	--	---	--

1.5. Report layout

The report is divided into 05 chapters as follows:

- Chapter 1. General introduction.

In the first chapter, a brief overview of the topic is presented. From there, objectives are set and solutions for the problem are proposed, as well as identifying the target audience and stakeholders involved with the topic.

- Chapter 2. Theoretical overview.

In this chapter, it provides a theoretical framework to understand the application of Enterprise Resource Planning (ERP) systems within specific project contexts. It begins with an introduction to ERP, outlining its fundamental principles, functionalities, and significance in business processes. Following this, the chapter delves into the core challenges and problems associated with implementing ERP solutions in real-world scenarios. To illustrate these issues practically, our study, in this project, which is recommendation system, is mentioned. The background knowledge as well as our approach are introduced.

- Chapter 3. Business process system.

In chapter 3, it explores the core components and processes of a business process system, providing an in-depth analysis of its structure and functionality. First the system overview offers insights into the overall architecture and integration of various business processes. Then, the subsequent subsections focus on specific processes, from procurement process to customer service process. These processes are modeled through Business Process Model and Notation (BPMN), outlines the specifications.

- Chapter 4. Proposed algorithms and solutions

In chapter 4, it discussed about how we implemented the research. This chapter clarifies the data collection process, how we do the data processing, algorithm explanation, and training model stage. Then, the practical results are given to evaluate the performance of the system.

- Chapter 5. Conclusions and future directions

This chapter gives a summary of the main findings of the research based on the result in chapter 4. It highlights the meaning of our project to the business and practical applications as well as suggests missions for the future development based on the objectives that is not completed.

1.6. Project plan – WBS

The Work Breakdown Structure (WBS) for the Brand Distribution Store Website Project is a detailed plan designed to guide the project from inception to launch. We distribute the project into many phrases for easy to work.

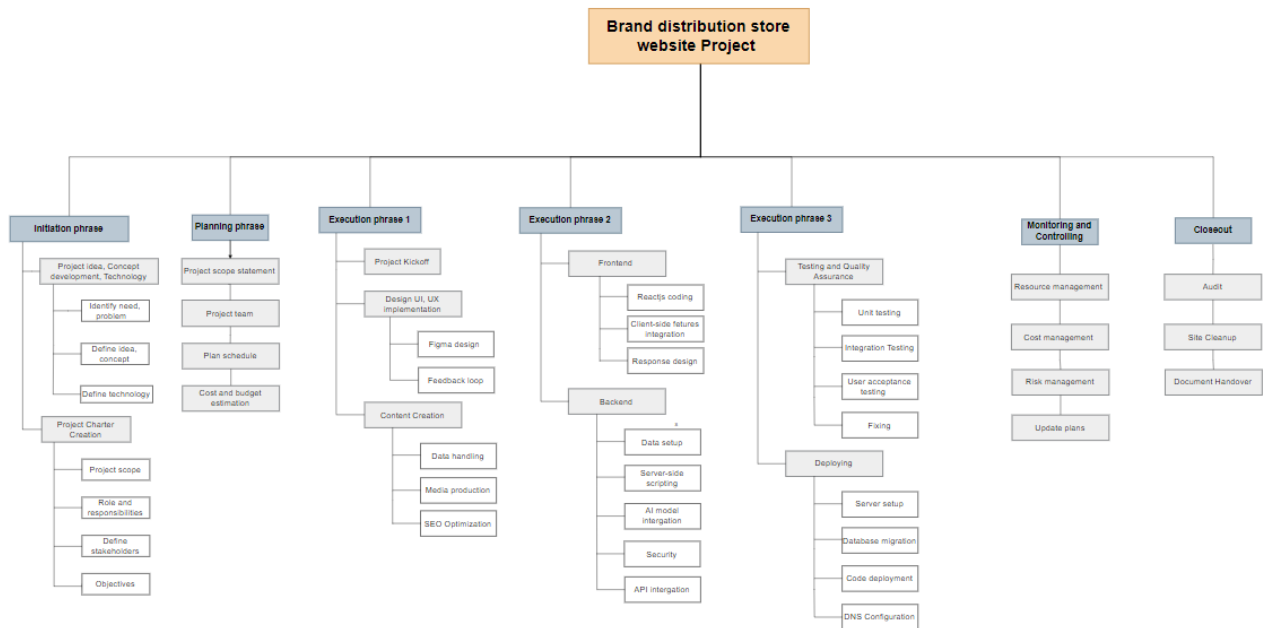


Figure 1: WBS

- **Project Initiation and Planning phase:** where the project scope is defined, a detailed plan and schedule are created, and roles are assigned.
- **The Execution phase:** Where the project has been created, because of the complication of implementation multiple technologies and function that why we decided to divide execution phase into 3 subparts.
 - **Execution phase 1:** The Design phase follows, involving requirement gathering, figma, and UI/UX design, data collection and analysis.
 - **Execution phase 2:** Building the front-end and back-end and integrating e-commerce functionalities.
 - **Execution phase 3:** Testing and deployment our platform.
- **Monitoring and Controlling:** involving overseeing project performance to ensure alignment with the plan, managing resources and costs, assessing, and mitigating risks, and updating plans as necessary to achieve project objectives.
- **Closeout:** conducting an audit to ensure all project requirements are met, performing site cleanup, and handing over all necessary documents to finalize the project.

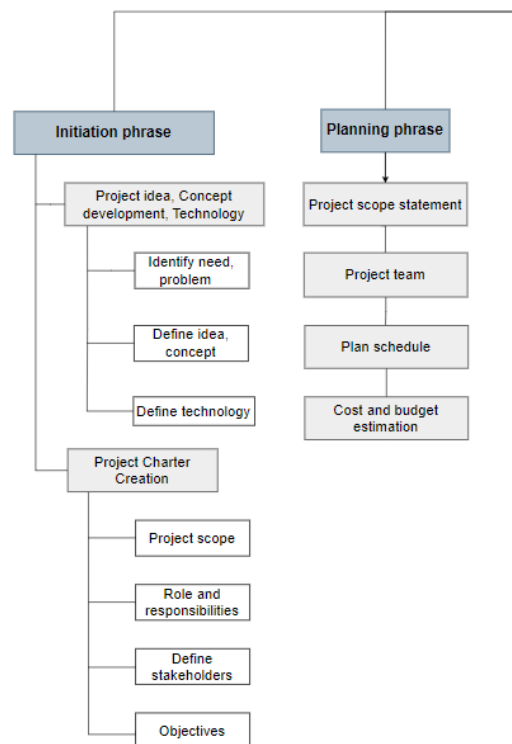


Figure 2: WBS - Initiation and Planning phases

- The initiation phase of a project is crucial as it sets the foundation and direction for all subsequent phases. This phase can be divided into two main components: Project Idea, Concept Development, and Technology, and Project Charter Creation.

1. Project Idea, Concept Development, and Technology

- Identify Need or Problem:

The initial step involves recognizing a need or problem that the project aims to address. This stems from market demands, organizational needs, or innovative opportunities.

- Define Idea or Concept:

Once the need or problem is identified, the next step is to brainstorm and outline the idea or concept that will address the issue. This involves creative thinking and feasibility analysis to ensure the idea is viable and worthwhile.

- Define Technology:

In this step, the necessary technology required to implement the idea is identified. Involving assessing existing technologies or developing new ones to ensure the project's success.

1. Project Charter Creation

- **Project Scope:**

The project scope outlines the boundaries of the project, specifying what will be included and what will not. It defines the work required to complete the project successfully, setting clear expectations and limits.

- **Role and Responsibilities:**

Defining the roles and responsibilities involves assigning specific tasks and duties to team members and stakeholders. This ensures that everyone knows their responsibilities and accountability, facilitating smooth project execution.

- **Define Stakeholders:**

Identifying and defining stakeholders involves determining all the individuals, groups, or organizations that will be affected by or can influence the project. Understanding stakeholder needs and expectations is crucial for project success.

- **Objectives:**

Setting clear and measurable objectives is vital to guide the project towards its desired outcomes. Objectives provide a benchmark for success and help keep the project focused and aligned with its goals.

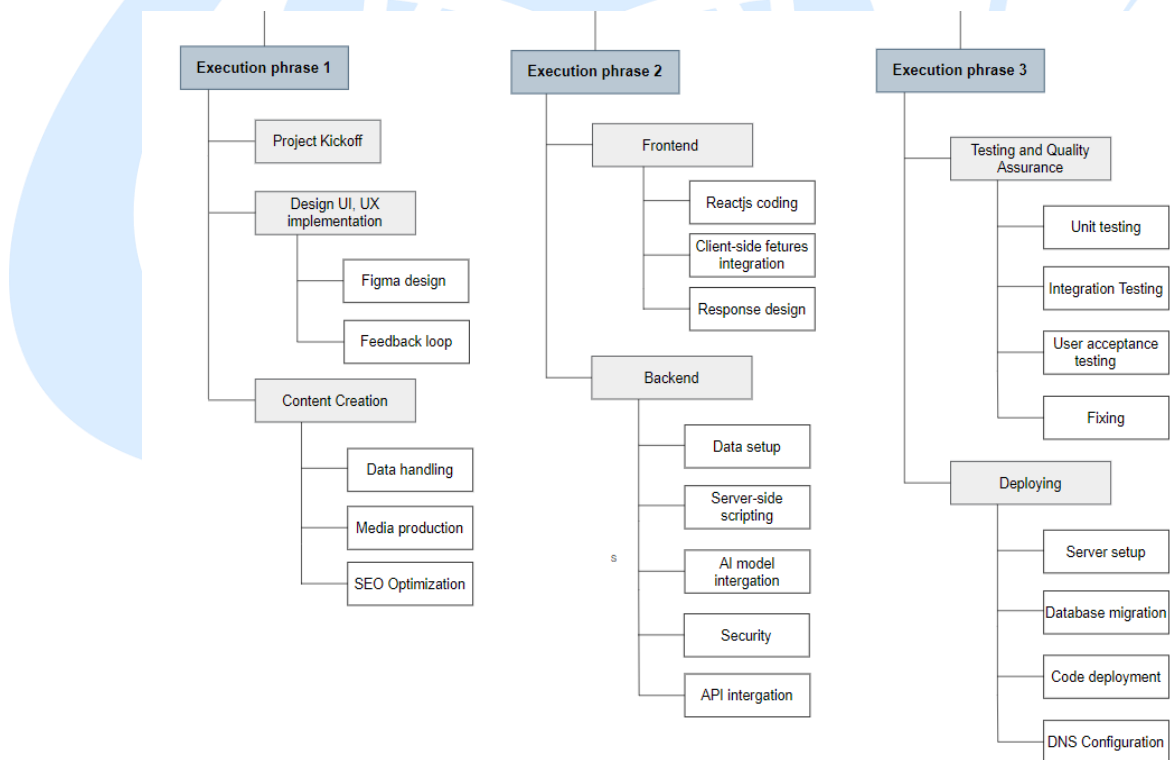


Figure 3: WBS - Execution phases

Execution Phase 1

This is the top-level component, indicating the first phase of the project execution. It's the overarching category under which all other tasks and sub-tasks fall.

1. **Project Kickoff:** The initial stage of the project where the team meets, outlines the project scope, objectives, and establishes roles and responsibilities. This is a crucial step to align the team's understanding and approach to the project.

2. **Design UI, UX Implementation:** This major task involves designing the user interface (UI) and user experience (UX). It is subdivided into:

- **Figma Design:**

Using Figma, a popular design tool, to create visual representations of the UI and UX. This involves wireframing, prototyping, and creating design mockups.

- **Feedback Loop:**

An iterative process where designs are reviewed, feedback is gathered, and revisions are made. This ensures the design meets user needs and stakeholder expectations.

3. **Content Creation:** Another significant task that encompasses creating various types of content required for the project. This is broken down into:

- **Data Handling:**

Managing and processing data required for the content. This could involve data collection, cleaning, and organization.

- **Media Production**

Creating media content such as images, videos, graphics, and other multimedia elements.

- **SEO Optimization**

Ensuring the content is optimized for search engines to enhance visibility and ranking. This involves keyword research, on-page SEO practices, and content adjustments to meet SEO standards.

Execution Phase 2

This is the second phase of the project execution, focusing on both the front-end and back-end development. It encapsulates the detailed tasks required to build and integrate the application's functionalities.

1. Frontend: This category includes all tasks related to the client-side development of the application, ensuring the user interface is functional and visually appealing. It is subdivided into:

- React.js Coding: Writing and maintaining the code using React.js, a popular JavaScript library for building user interfaces. This involves creating components, managing states, and handling events.
- Client-side Features Integration: Implementing and integrating various features that will run on the client-side of the application. This might include forms, interactive elements, and other front-end functionalities.
- Responsive Design: Ensuring the application is responsive and provides an optimal viewing experience across various devices and screen sizes. This involves using CSS and other design techniques to make the UI adaptable.

2. Backend

- This category includes all tasks related to the server-side development of the application, ensuring the backend logic and database interactions are properly handled. It is subdivided into:
 - Data Setup: Setting up the database and organizing the data structures required for the application. This involves database design, creating tables, and defining relationships.
 - Server-side Scripting: Writing server-side scripts to handle the business logic, data processing, and interactions with the frontend. This can include creating RESTful APIs and managing server operations.
 - AI Model Integration: This involves using the AI pre-trained models, embedding them into the backend, and ensuring they interact seamlessly with the rest of the system in semantic search, and recommender system.
 - Security: Implementing security measures to protect the application from threats and vulnerabilities. This includes authentication, authorization, data encryption.

- **API Integration:** Connecting the application to external services and APIs to extend its functionality. This involves consuming third-party APIs, handling requests and responses, and ensuring smooth integration.

Execution Phase 3

This is the third phase of the project execution, testing and deploying our project. In this phase we are going to apply unit testing, integration testing, user acceptance testing and fixing the unreasonable point. We also use AWS cloud in order to deploy the project.

1. **Testing and Quality Assurance:** Testing and Quality Assurance involves systematically evaluating and verifying that a project's deliverables meet the specified requirements and standards, ensuring functionality, reliability, and overall quality.

- **Unit testing:** Unit testing involves testing individual components or modules of the software to ensure they function correctly in isolation. This helps catch bugs early and ensures each part performs as expected.
- **Integration Testing:** Integration testing focuses on verifying the interactions between different components or systems. This ensures that combined parts work together as intended, identifying issues in the interfaces and interaction logic.
- **User acceptance testing:** User acceptance testing is conducted by the end-users or clients to validate that the software meets their requirements and performs its intended functions in a real-world scenario. UAT is crucial for gaining user approval before final deployment.
- **Fixing:** Fixing involves addressing any defects or issues identified during the testing phases. This step is essential for refining the software and ensuring it meets all quality standards and performs reliably.

2. **Deploying:** Using AWS cloud service for deploying.

- **Setup server:** Setting up the server includes provisioning the necessary AWS infrastructure, EC2 instances.
- **Database migration:** Using RDS service for deploying our relational database.
- **Code playing transferring the application code to the AWS server.** This includes setting up deployment pipelines using services like AWS Code Deploy, ensuring the application is correctly installed and configured.

- **DNS configuration:** DNS configuration involves setting up the Domain Name System (DNS) to ensure that the application is accessible via a user-friendly domain name. This includes configuring Route 53 or other DNS services to map the domain name to the application's IP address.

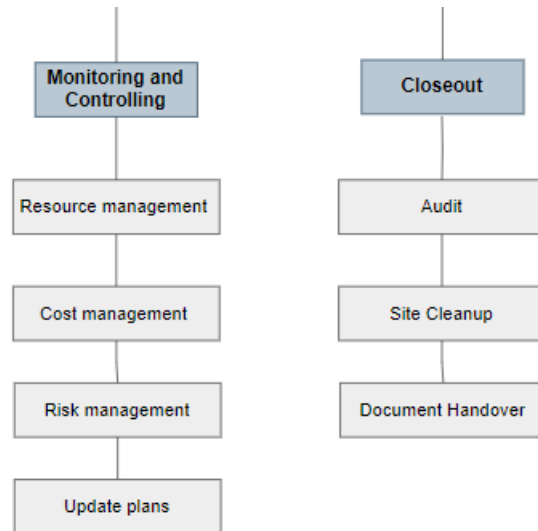


Figure 4: WBS - Monitoring and Closeout phases

Monitoring and Controlling

This is a critical phase in project management that involves overseeing all project activities to ensure they align with the project plan. It encompasses tracking progress, identifying issues, and adjusting as necessary to keep the project on track.

1. Resource Management

This task involves managing the project's resources effectively. It includes allocating resources (such as personnel, equipment, and materials) as needed, ensuring optimal use, and adjusting allocations to meet changing project demands.

2. Cost Management

This task focuses on monitoring and controlling project costs. It involves tracking expenditures, comparing actual costs against the budget, and making financial adjustments to avoid cost overruns. This ensures the project stays within the approved budget.

3. Risk Management

This task entails identifying, analyzing, and responding to project risks. It involves continuous risk assessment to detect potential issues early, implementing mitigation strategies, and managing responses to any risks that materialize to minimize their impact on the project.

4. Update Plans

This task involves revising project plans based on monitoring and controlling activities. It includes updating schedules, adjusting resource plans, refining cost estimates, and incorporating changes in scope or deliverables. This ensures the project plan remains relevant and achievable as the project progresses.

Closeout: The closeout phase involves ensuring that all project activities are completed, deliverables are handed over, and the project is formally closed.

1. **Audit:** An audit involves a thorough review of the project to ensure that all objectives were met, deliverables were completed to the required standards, and all processes were followed correctly. This helps identify lessons learned and areas for improvement in future projects.

2. **Site cleanup:** Site cleanup entails clearing and organizing the physical or digital project site. This step ensures that any temporary setups, equipment, or data used during the project are properly disposed of or archived, leaving the site in good condition.

3. **Document handover:** Document handover involves transferring all relevant project documentation to the client or relevant stakeholders. This includes final reports, user manuals, project plans, and any other documentation that will help maintain and support the project deliverables in the future.

1.7. Work assignment table

Table 2: Work assignment

PHRASE	TASK	SUB TASK	Dang Thanh Dat	Ngo Thuan Phat	Nguyen Lam Nhat Anh	Vo Thi Thu Tien	Date	Status
--------	------	----------	----------------------	----------------------	------------------------------	--------------------------	------	--------

Initiation phrase	Project idea, Concept development, Technology	Identify need, problem	✓	✓	✓	✓	14/4/2024 - 20/4/21024	Done
		Define idea, concept	✓	✓	✓	✓	14/4/2024 - 20/4/21025	Done
		Define technology	✓	✓	✓	✓	14/4/2024 - 20/4/21026	Done
	Project Charter Creation	Project scope	✓	✓	✓	✓	14/4/2024 - 20/4/21027	Done
		Role and responsibilities	✓	✓	✓	✓	14/4/2024 - 20/4/21028	Done
		Define stakeholders	✓	✓	✓	✓	14/4/2024 - 20/4/21029	Done
		Objectives	✓	✓	✓	✓	14/4/2024 - 20/4/21030	Done
	Project scope statement		✓	✓	✓	✓	20/4/2024-	Done

Planning phrase							24/4/2024	
	Project team		✓	✓	✓	✓	20/4/2024-24/4/2025	Done
	Plan schedule				✓	✓	20/4/2024-24/4/2026	Done
	Cost and budget estimation		✓	✓	✓	✓	20/4/2024-24/4/2024	Done
Execution phrase 1	Project Kickoff		✓	✓	✓	✓	24/4/2024-30/4/2024	Done
	Design UI, UX implementation	Figma design		✓		✓	24/4/2024-30/4/2025	Done
		Feedback loop		✓		✓	24/4/2024	In Progress
	Content Creation	Data handling	✓				30/4/2024 - 4/5/2024	Done
		Media production			✓		30/4/2024 - 4/5/2025	Done

		SEO Optimization	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	30/4/2024 - 4/5/2024	In Progress
Execution phrase 2	Frontend	Reactjs coding	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	4/5/2024-30/5/2024	In Progress
		Client-side fetures integration	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	4/5/2024-30/5/2025	In Progress
		Response design	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	4/5/2024-30/5/2026	In Progress
	Backend	Data setup	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	7/5/2024-10/5/2024	Done
		Server-side scripting	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	7/5/2024-10/5/2025	Done
		AI model intergation	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10/5/2024 - 30/5/2024	In Progress
		Security	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	10/5/2024-15/5/2024	Done
		API intergation	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	10/5/2024 - 30/5/2024	In Progress

Execution phrase 3	Testing and Quality Assurance						10/5/2024 - 30/5/2024	In Progress
		Unit testing			✓	✓		
		Integration Testing	✓	✓			10/5/2024 - 30/5/2025	Stuck
		User acceptance testing	✓	✓	✓	✓	10/5/2024 - 30/5/2024	In Progress
		Fixing	✓	✓	✓	✓	10/5/2024 - 30/5/2024	In Progress
	Deploying	Server setup			✓		15/5/2024 - 20/5/2024	Done
		Database migration	✓		✓		15/5/2024 - 20/5/2025	Done
		Code deployment	✓		✓		10/5/2024 - 30/5/2024	In Progress
		DNS Configuration	✓		✓		15/5/2024 -	Stuck

							20/5/2024	
Monitoring and Controlling	Resource management		<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	25/5/2024-30/5/2024	In Progress
	Cost management		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	25/5/2024-30/5/2025	In Progress
	Risk management		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	25/5/2024-30/5/2026	In Progress
	Update plans		<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	25/5/2024-30/5/2027	In Progress
Closeout	Audit		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30/5/2024-4/6/2024	In Progress
	Site Cleanup		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30/5/2024-4/6/2025	Stuck
	Document Handover		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	30/5/2024-4/6/2026	In Progress

CHAPTER 2: THEORETICAL OVERVIEW

In this chapter, it provides a theoretical framework to understand the application of Enterprise Resource Planning (ERP) systems within specific project contexts. It begins with an introduction to ERP, outlining its fundamental principles, functionalities, and significance in business processes. Following this, the chapter delves into the core challenges and problems associated with implementing ERP solutions in real-world scenarios. To illustrate these issues practically, our study, in this project, which is recommendation system, is mentioned. The background knowledge as well as our approach are introduced.

2.1. ERP introduction

Enterprise Resource Planning (ERP) systems have emerged as a cornerstone in the contemporary business landscape, particularly vital for managing and integrating various functions across an organization. ERP systems are software platforms designed to automate core business processes, including finance, human resources, manufacturing, supply chain, services, procurement, and others. By consolidating these functions into a unified system, ERP solutions enhance efficiency, accuracy, and productivity, providing businesses with a holistic view of their operations.

In today's fast-paced and highly competitive market, the role of ERP systems has become more crucial than ever. The integration capabilities of ERP systems allow organizations to break down silos and foster collaboration across departments, leading to better decision-making and strategic planning. This integration is particularly beneficial for retail businesses, such as clothing stores, where managing inventory, sales, procurement, and customer relations seamlessly is critical to maintaining a competitive edge.

The Role of ERP in the Modern Era

The digital transformation wave has swept across industries, and ERP systems are at the forefront of this transformation. In the era of Big Data, the Internet of Things (IoT), and Artificial Intelligence (AI), ERP systems are evolving to incorporate advanced technologies that further enhance their functionality and value. These systems are no longer just about

automating back-office functions; they are now central to driving business intelligence, improving customer experiences, and enabling agile responses to market changes.

For a clothing store, an ERP system can transform operations in several impactful ways. Inventory management, for example, becomes more efficient with real-time tracking and automated reordering processes, ensuring that popular items are always in stock and reducing the risk of overstocking less popular products. Sales data can be analyzed to forecast trends and adjust purchasing strategies accordingly. Moreover, integrating customer relationship management (CRM) within the ERP system can provide valuable insights into customer preferences and buying behaviors, allowing the store to tailor its marketing efforts and improve customer satisfaction.

2.2. Technologies used in system building

2.1.1. Overview of technology

To build and support the Brand Distributor Store System, our development team will utilize modern and efficient tools and technologies known for their robustness and scalability:

- **Git / GitHub:** For version control and collaborative development, ensuring that all code changes are tracked and managed efficiently.
- **Visual Studio Code:** As the primary code editor due to its versatility and support for multiple programming languages and tools.
- **React:** Employed as the front-end framework to build a dynamic and responsive user interface.
- **MySQL Database:** Used as the backend database system to handle large volumes of data with high performance and reliability.
- **Node.js /Express:** For the server-side logic, providing a lightweight and efficient platform for handling web requests.

These tools and technologies were selected to optimize the development process and ensure the system is scalable, maintainable, and meets modern web standards.

2.1.2. History of development

Today, technical functions supporting for the customer plays a vital role in the survival and development of organizations and businesses especially in e-commerce field.

Increasingly, there are many and many functions were born to enhance the shopping experience and reap the huge profits from customers. Among the a number of inovations, the ability to find the exacly product based on the specific customers' demands and recommend protential products are contributing a significant profits for e-commerce organizations.

Regarding the former, finding the product that highly releted to the customers' demend have thrown a lot of challanging. Using searching method based on the SQL platform required the customers to type the exacly words on the search bar, which make them fell uncomfortable and may be take a lot of time for remmenbering and typing in applications. Other platform is elastic search compared to SQL, offers powerful full-text search capabilities and scalability for large datasets. While SQL excels at structured data queries and relational operations, Elasticsearch is optimized for real-time search and analytics, handling unstructured data efficiently. Elasticsearch's distributed nature ensures high availability and fast search performance, making it ideal for applications requiring rapid and complex search functionalities. But these methods still not satisfy the demand for searching by description of the product or customers' memories. According to these limitations of previous methods and the actual needs of customers our team propose the semantic method call Hyper search based on the combination of both features of product image and text data including product name, brand name, product description. By using the deep learning and AI method we can understand the customers' need and then give them the products that most related to their query. Semantic search, compared to Elasticsearch and SQL, focuses on understanding the context and intent behind queries to deliver more relevant results. This makes semantic search particularly powerful for applications requiring nuanced understanding and contextual relevance.

In addition, recommendations system has been divided into three main model populations filtering, content-based filtering, collaborative filtering which required a huge database which push enterprises under the pressure of collecting data and required a significant resource for training and maintenance. But in the recent scientific literature, there is a new approach named vector database CLIP model of OpenAI which reduce the pressure of data and training cost, but the method can improve the accuracy by using the personal data of the customer such as searching and order history to understand the habits and

preferences during the time that they use our application. That is why we choose vector database approach for our recommendations system and expect for the positive contribution to our project.

2.3. Description of the problem applying ERP to the project topic (or Case study)

2.3.1. Hybrid search using image vision and semantic of text

Nowadays, multiple search engines on many e-commerce platforms serve searches based on product names or descriptions. This presents challenges for customers during the buying process for three primary reasons:

Firstly, product descriptions may not be relevant to the product images, leading to customer disappointment and discouraging return visits to the platform. Secondly, these search engines often fail to meet customer demands, especially when customers search based on memories or partial descriptions of the product. Lastly, the search speed of traditional search engines can waste customers' time, negatively impacting their overall experience.

These issues are concerning and highlight the need for a solution to improve customer experience and assist ERP systems in upgrading their sales processes. According to our research we tried to develop CLIP model represents a significant advancement in semantic search capabilities, integrating image vision and language understanding to facilitate more intuitive and efficient data retrieval. When applied to ERP (Enterprise Resource Planning) systems, CLIP can bring numerous benefits to businesses, and can handle these problems above.

Firstly, the CLIP model offers a powerful solution to the problem of irrelevant product descriptions by merging the spaces of product images and text descriptions. This model effectively aligns visual and textual representations, enabling it to understand and correlate images with their corresponding descriptions accurately. By integrating CLIP into an ERP system, businesses can ensure that product images and descriptions are consistently relevant and aligned, significantly enhancing customer satisfaction. This integration not only improves the accuracy of search results but also increases customer trust and engagement, leading to more efficient sales processes and better overall performance of the ERP system.

One of the primary advantages of integrating CLIP into ERP systems is the enhancement of data retrieval processes. Traditional ERP systems often require precise keywords or specific queries to locate information, which can be time-consuming and inefficient. CLIP, however, allows users to employ natural language queries, which are matched semantically with relevant data. This means that users can describe what they are looking for in everyday language, and the system will understand and retrieve the most relevant information, whether it is text, documents, or images. This significantly reduces the time and effort required to find information, leading to increased productivity.

Improved efficiency and productivity translate into cost savings. Reducing the time employees spend searching for information allows them to focus on more value-added activities. Additionally, the increased accuracy of search results minimizes errors and the associated costs of rectifying them. Over time, these savings can be substantial, contributing to a healthier bottom line.

In conclusion, the CLIP model significantly enhances e-commerce platforms by merging product image and text spaces, ensuring accurate and relevant search results. This improvement leads to higher customer satisfaction, increased trust, and more efficient sales processes, ultimately boosting the overall performance of e-commerce systems.

2.3.2. Case study

Challenge:

Trendy Threads relied on a traditional keyword-based search system. Customers often struggled to describe desired clothing items accurately using text alone, leading to frustration and missed sales opportunities. For example, a customer looking for a "flowy floral dress with a ruffled neckline" might get irrelevant results for "floral dress" or miss the perfect dress simply because they didn't use the exact term "ruffled neckline."

Solution:

Trendy Threads implemented a CLIP-powered search engine. They gathered a massive dataset of product images and corresponding detailed descriptions. This data was used to

train a CLIP model, enabling it to understand the relationships between the visual features of clothing (e.g., patterns, cuts, styles) and their textual descriptions.

Results:

Effortless Visual Search: CLIP analyzes the image and retrieves similar products from Trendy Threads' vast inventory. This empowers customers to find specific styles, patterns, or unique details they might struggle to describe in text.

Enhanced Text Search: CLIP refines traditional text-based search. Even with vague queries like "summer dress," CLIP analyzes both product descriptions and images to suggest a wider range of relevant options. This is particularly helpful for customers who have a general idea but lack specific keywords.

Increased Sales: Since implementing CLIP, Trendy Threads has seen a significant increase in conversion rates. Customers can now find the exact items they're looking for more easily, leading to more purchases.

Improved Customer Satisfaction: Faster and more accurate product discovery translates to a more enjoyable shopping experience. Customers are more likely to return to Trendy Threads for future fashion finds.

Conclusion:

Trendy Threads' adoption of CLIP demonstrates the transformative power of this technology in e-commerce. By bridging the gap between a customer's visual intent and textual product descriptions, CLIP empowers a smoother, more intuitive search experience. This translates to increased sales, happier customers, and a thriving online fashion destination.

CHAPTER 3: BUSINESS PROCESS SYSTEM

In this chapter, we focus on Business Process Model and Notation (BPMN), examining its specifications and applications. We explore how BPMN serves as a standard for business process modeling, providing a graphical representation that is easy to understand and implement across various business functions. This discussion will also cover key elements and their roles in enhancing communication and efficiency in business processes. Through this exploration, readers will gain insights into effectively utilizing BPMN to optimize their organizational operations.

3.1. System overview

The system overview delves into the architecture and functionalities of Business Process Model and Notation (BPMN) within the framework of enterprise operations. BPMN is designed to bridge the gap between business process design and process implementation, offering a standardized method to describe and document processes in a format that is both technically precise and accessible to all stakeholders.

This section outlines the core components of BPMN, including process diagrams, decision points, tasks, and flow sequences. It discusses how these elements are used to map out complex business activities in a clear and methodical way, allowing for effective analysis and continuous improvement. Additionally, the overview examines the role of BPMN in facilitating the integration of new technologies into business processes, such as automation tools and data analytics platforms, which can lead to significant enhancements in productivity and decision-making.

By providing a thorough understanding of BPMN's structure and capabilities, this section equips readers with the knowledge to assess and refine their business processes, ensuring they align with strategic objectives and adapt efficiently to changing market demands.

[illegible]

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
----------	---------	-----------	-------------

Identify Product		Procurement Department	Staff identifies a product that needs to be purchased
Create Purchase Request		Procurement Department	Staff creates a purchase request
Send Request		Procurement Department	The request is sent for approval
Approve	Exclusive Gateway		Manager approves or rejects the request
Send Deposit		Financial Accounting Department	If approved, a deposit is sent
Notification Sent		Procurement Department	Notification is sent to the supplier
Request Received		Supplier	Supplier receives the request
Check Request		Supplier	Supplier checks the request
Select Supplier	Exclusive Gateway	Procurement Department	If the supplier is known, they are selected. If not, research is required
Research Supplier		Procurement Department	Research is conducted to find a suitable supplier
Create RFP/RFQ		Procurement Department	A request for proposal (RFP) or request for quotation (RFQ) is created
Send RFP/RFQ		Procurement Department	The RFP/RFQ is sent to potential suppliers
Supplier Receives RFP/RFQ		Supplier	Supplier receives the RFP/RFQ

Review Bid		Procurement Department	Procurement department reviews bids from suppliers
Supplier submits bid		Supplier	Supplier submits a bid
Bid Received		Procurement Department	The bid is received from the supplier
Select Bidder		Procurement Department	The best bid is selected
Contract/Order		Procurement Department	A contract or purchase order is created
Review Contract/Order		Supplier	Supplier reviews the contract/order
Rejection Received	Exclusive Gateway	Procurement Department	If the supplier rejects the contract/order, the process ends with a rejection letter sent
Confirmation Received	Exclusive Gateway	Procurement Department	If the supplier confirms the contract/order, the process moves forward
Complete Order Procedure		Procurement Department	The procurement department finalizes the order
Notification Sent		Procurement Department	Notification is sent to other departments (e.g., Inventory)
Receive Notification		Inventory Department	Inventory department receives notification
Check Products		Inventory Department	Staff checks the quality and quantity of received products
Update Inventory		Inventory Department	Inventory records are updated

3.3. Sales process

3.3.1. BPMN

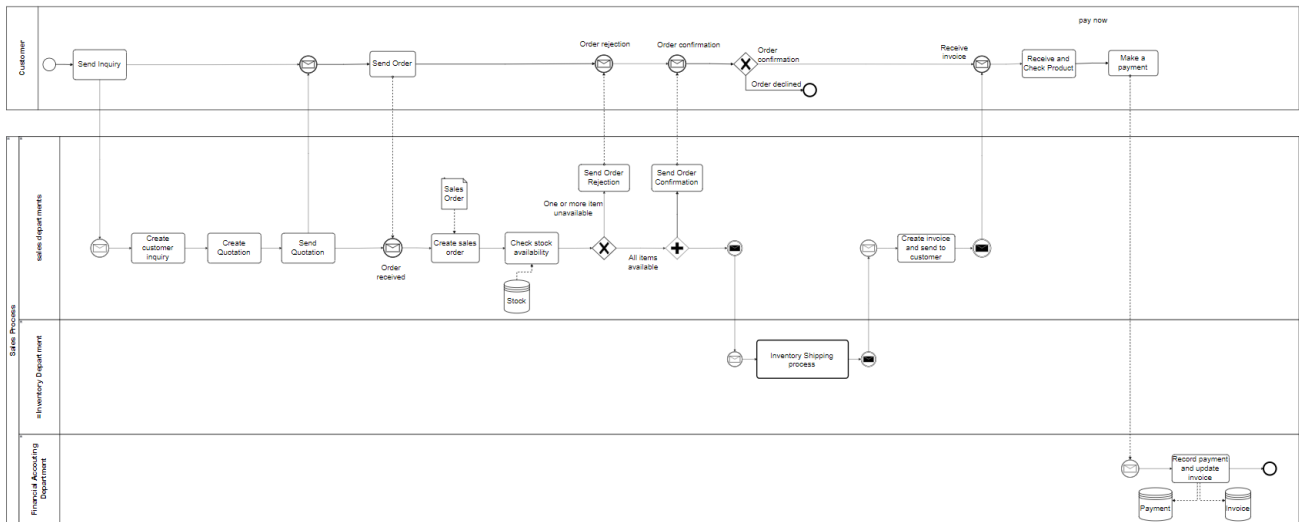


Figure 6: BPMN - Sale process

3.3.2. Specification

Table 4: Sale process specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Sand Impurity Test		Inventory Department	Inventory department tests the sand for impurities
Inventory Check		Inventory Department	Inventory department checks stock levels
Inventory Sufficient (Yes)	Exclusive Gateway	Inventory Department	There is enough sand in stock to proceed
Inventory Sufficient (No)	Exclusive Gateway	Inventory Department	There is not enough sand in stock
Update Inventory		Inventory Department	Inventory records are updated
Order Sand		Purchasing Department	A purchase order is created to order more sand

Sand Received		Inventory Department	The inventory department receives the sand order
Sign Off			A supervisor or designated role signs off on the process

3.4. Inventory receiving process

3.4.1. BPMN

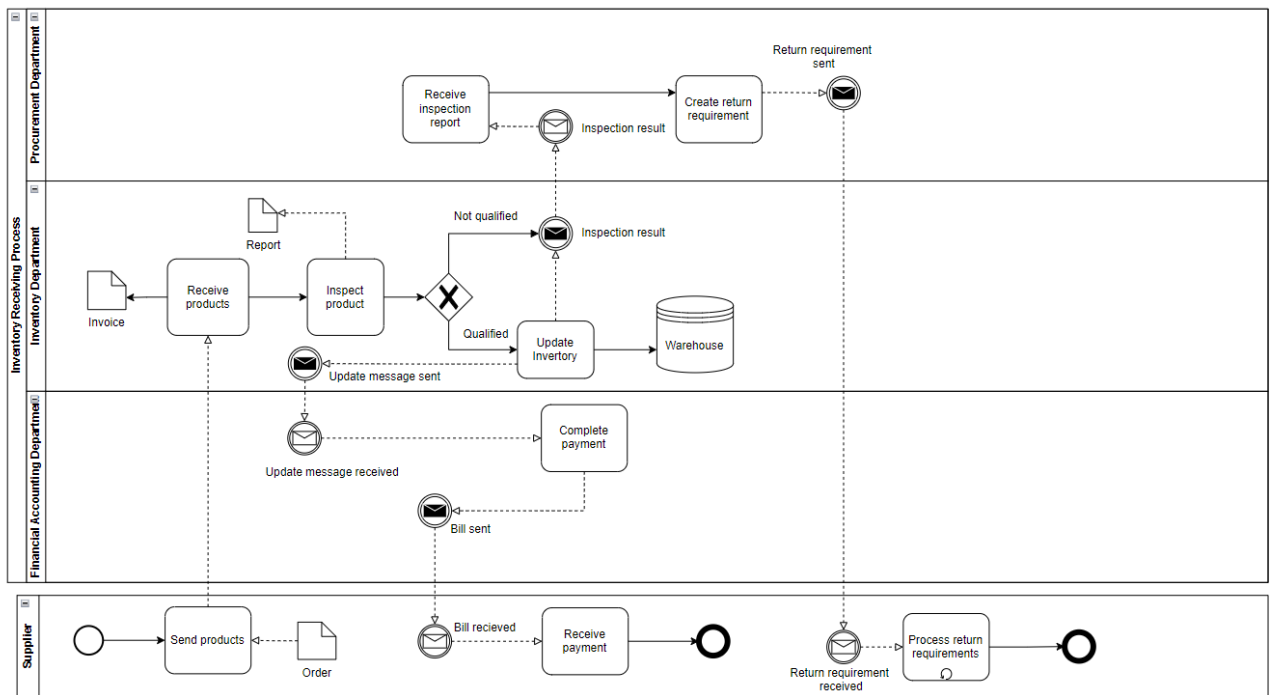


Figure 7: BPMN - Inventory receiving process

3.4.2. Specification

Table 5: Inventory receiving process specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Receive Products		Inventory Receiving Process	Products are received from the supplier
Receive Invoice		Inventory Receiving Process	Invoice is received from the supplier
Receive Return Requirement		Inventory Receiving Process	A request to return products may be received

Inspect Products		Inventory Receiving Process	Products are inspected to verify quality and quantity
Update Inventory		Inventory Department	Inventory records are updated based on the received products
	Exclusive Gateway	Inventory Receiving Process	
Qualified		Inventory Receiving Process	Products meet quality standards
Create Return Requirement		Inventory Receiving Process	If products do not meet quality standards, a return request is created
Send Return Requirement		Inventory Receiving Process	The return request is sent to the supplier
Not Qualified		Inventory Receiving Process	Products do not meet quality standards
Send Bill to Accounting		Inventory Receiving Process	The invoice is sent to the accounting department
Process Return Requirements		Inventory Receiving Process	The return request is processed (outside the scope of this diagram)
Complete Payment		Financial Accounting Department	Payment is sent to the supplier
Bill Received		Inventory Receiving Process	The accounting department receives the bill

3.5. Inventory shipping process

3.5.1. BPMN

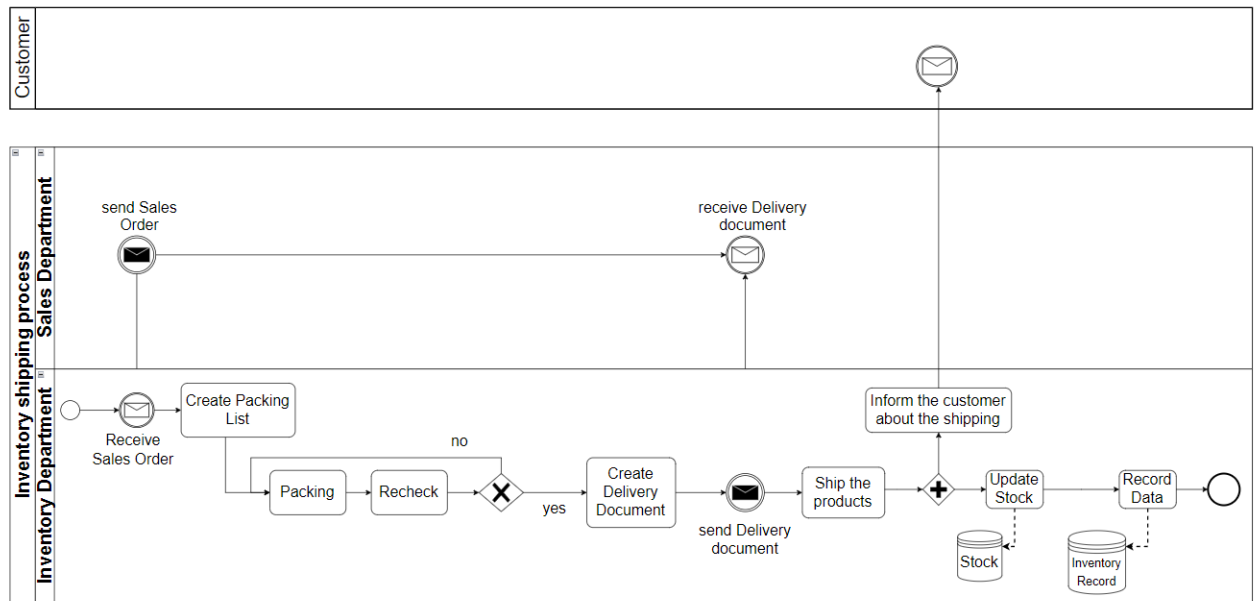


Figure 8: Inventory shipping process

3.5.2. Specification

Table 6: Inventory shipping process specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Receive Sales Order		Inventory Department	The inventory department receives a sales order
Check Inventory Levels		Inventory Department	Inventory staff checks to see if sufficient inventory is available to fulfill the order
Inventory Sufficient (Yes)	Exclusive Gateway	Inventory Department	There is enough inventory to fulfill the order
Inventory Sufficient (No)	Exclusive Gateway	Inventory Department	There is not enough inventory to fulfill the order
Create Picking List		Inventory Department	A picking list is created to specify the items required to fulfill the order
Package Order		Warehouse	The order is picked, packaged and labelled for shipment
Ship Order		Warehouse	The order is shipped to the customer

Update Inventory Records		Inventory Department	Inventory records are updated to reflect the outgoing order
Send Invoice		Sales Department	The sales department sends an invoice to the customer
Place Purchase Order		Inventory Department	A purchase order is placed to replenish inventory levels
Receive Notification		Inventory Department	The inventory department receives notification that the order has been shipped
Inform Customer		Sales Department	The sales department informs the customer that the order has been shipped

3.6. Financial and Accounting management process

3.6.1. BPMN

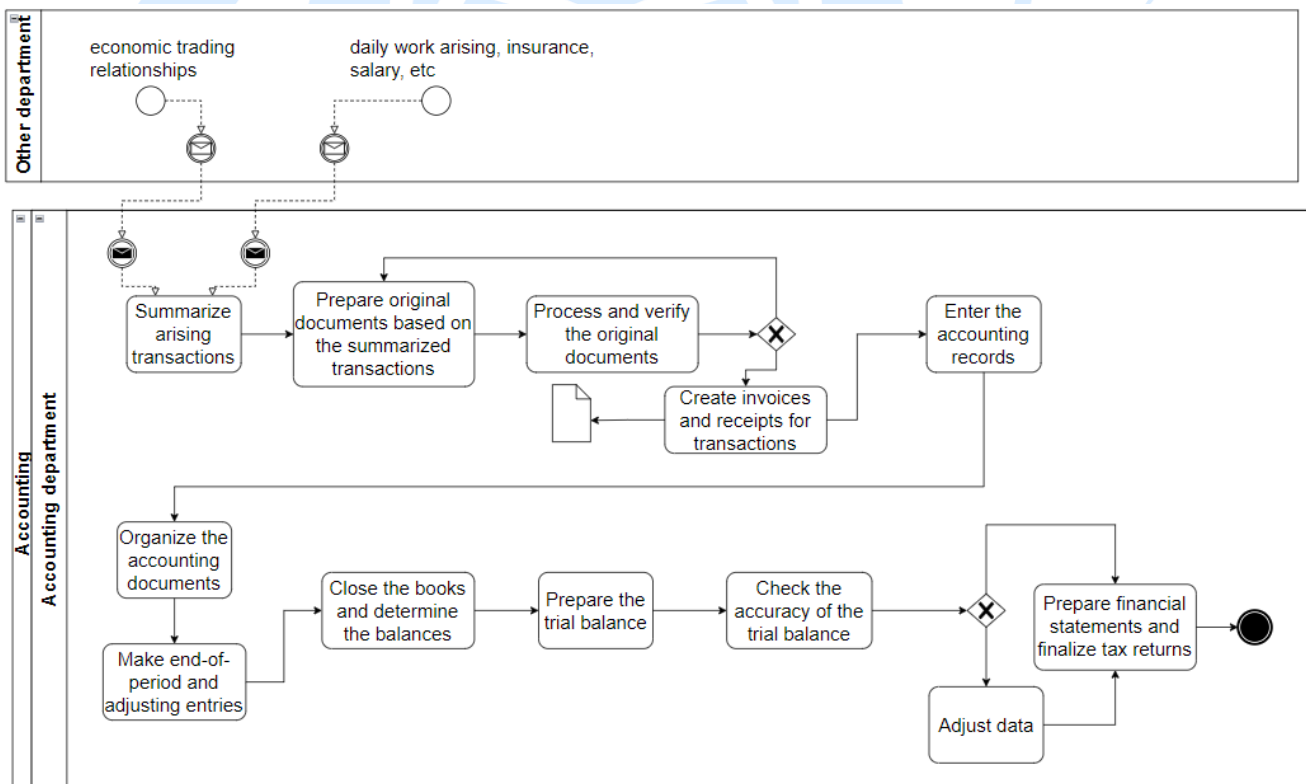


Figure 9: BPMN - Financial and Accounting management process

3.6.2. Specification

Table 7: Human resources management process specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Summarize Transactions		Accounting Department	Accounting staff summarizes the daily accounting transactions
Prepare Original Documents		Accounting Department	Staff creates original documents based on the summarized transactions
Process and Verify Original Documents		Accounting Department	The documents are processed to ensure accuracy
Enter Accounting Records		Accounting Department	The accounting records are entered into the accounting system
Organize Accounting Documents		Accounting Department	The accounting documents are organized for future reference
Close the Books	Exclusive Gateway	Accounting Department	The books are closed for the accounting period
Close the Books (Yes)		Accounting Department	The books are closed
Determine Balances		Accounting Department	The account balances are determined
Prepare Trial Balance		Accounting Department	A trial balance is prepared to verify the accuracy of the accounting records
Check Accuracy of Trial Balance		Accounting Department	The trial balance is checked for accuracy
Trial Balance Accurate (Yes)	Exclusive Gateway	Accounting Department	The trial balance is accurate

Trial Balance Accurate (No)	Exclusive Gateway	Accounting Department	The trial balance is not accurate
Make Adjusting Entries		Accounting Department	Adjusting entries are made to ensure the accuracy of the financial statements
Prepare Financial Statements		Accounting Department	Financial statements, such as the income statement and balance sheet, are prepared
Finalize Tax Returns		Accounting Department	Tax returns are finalized
Adjust Data		Accounting Department	Accounting data may be adjusted based on the finalization of tax returns

3.7. Human resources management process

3.7.1. BPMN

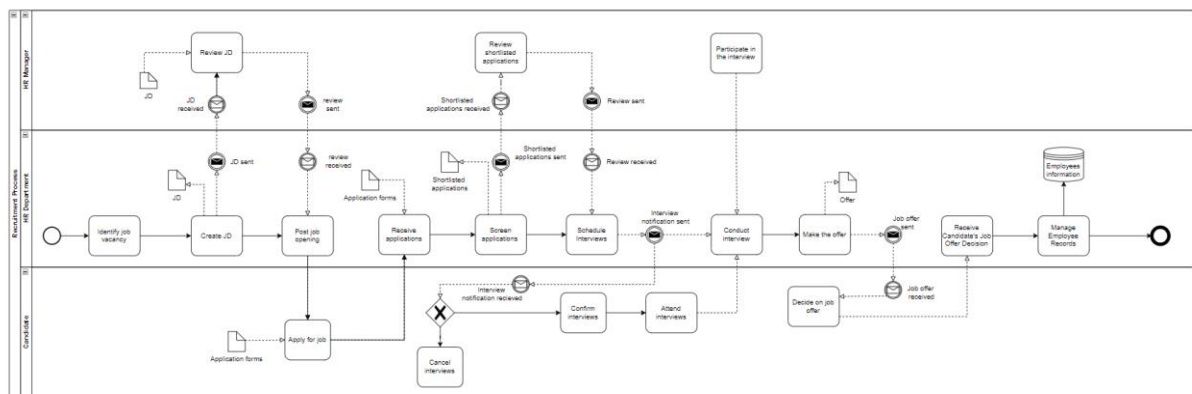


Figure 10: BPMN - Recruitment process

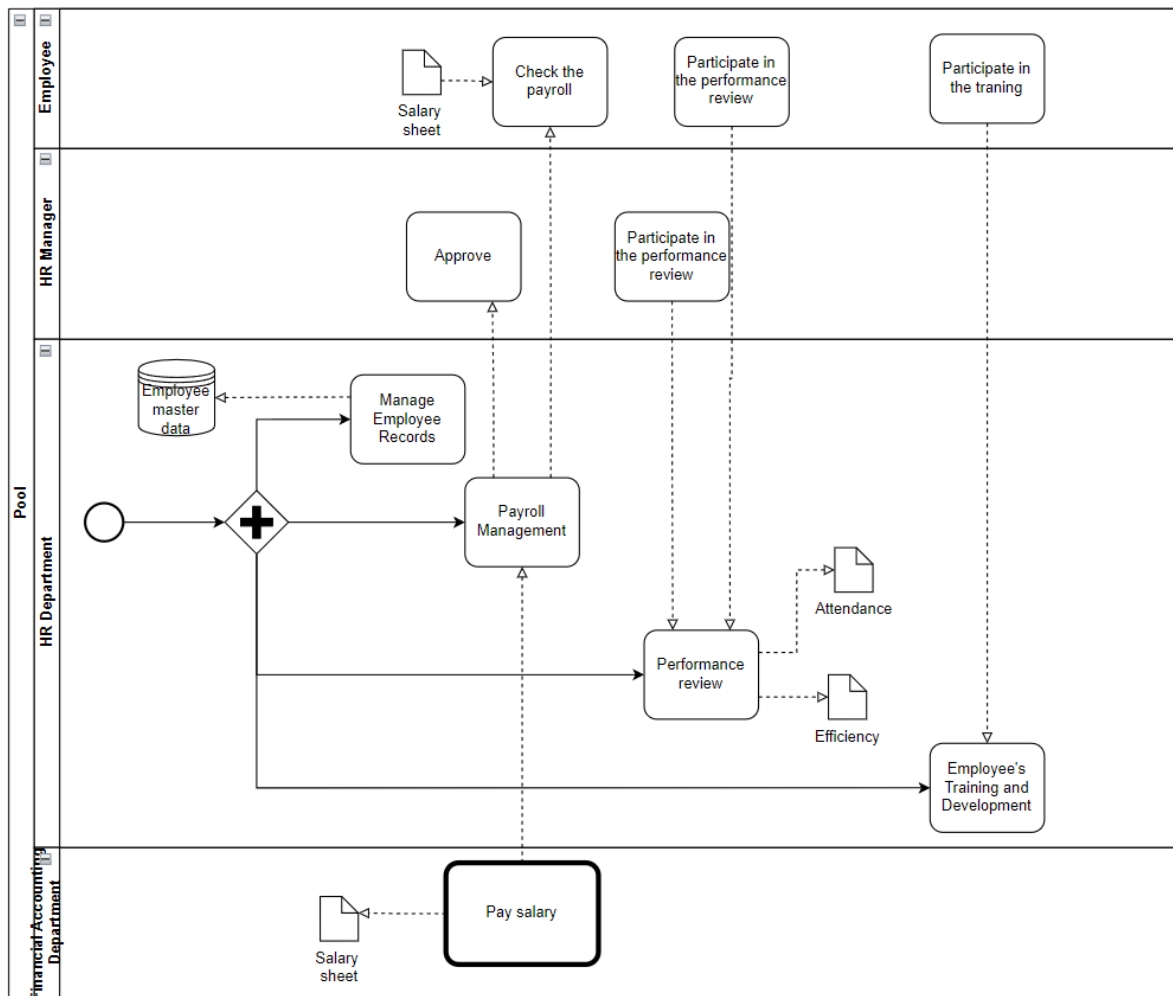


Figure 11: Human resources management process

3.7.2. Specification

Table 8: Recruitment process

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Apply for a Job		Candidate	Candidate applies for a job
Receive Application		Recruiting Department	Recruiting department receives the application
Attend Pre-screening		Candidate	Candidate attends a pre-screening interview
Pre-screening Interview		Recruiting Department	Recruiting department conducts a pre-screening interview

Screened In (Yes)	Exclusive Gateway	Recruiting Department	Candidate is selected to move forward in the process
Screened In (No)	Exclusive Gateway	Recruiting Department	Candidate is not selected to move forward in the process
Send Rejection Letter		Recruiting Department	A rejection letter is sent to the candidate
Conduct Candidate Evaluation		Hiring Manager	The hiring manager evaluates the candidate
Schedule Interview		Recruiting Department	If the candidate is chosen for an interview, an interview is scheduled
Interview		Hiring Manager	The hiring manager interviews the candidate
Interview Outcome (Pass)	Exclusive Gateway	Hiring Manager	Candidate passes the interview
Interview Outcome (Fail)	Exclusive Gateway	Hiring Manager	Candidate fails the interview
Send Rejection Letter (2nd Time)		Recruiting Department	A rejection letter is sent to the candidate
Offer Job		Hiring Manager	The hiring manager offers the job to the candidate
Make Job Decision		Candidate	The candidate decides to accept or reject the job offer
Accept Job Offer	Exclusive Gateway	Candidate	Candidate accepts the job offer
Reject Job Offer	Exclusive Gateway	Candidate	Candidate rejects the job offer
Send Offer Letter		Recruiting Department	An offer letter is sent to the candidate

Receive Signed Offer Letter		Recruiting Department	The recruiting department receives the signed offer letter
Complete Onboarding Process		Human Resources	The HR department completes the onboarding process for the new hire

Table 9: Human resources management process specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Check Payroll		HR Department	Human Resources department checks payroll data
Participate in Performance Review		Employee	Employee participates in a performance review
Participate in Training		Employee	Employee participates in company training
Approve Salary	Exclusive Gateway	Manager	Manager approves or rejects employee salary
Manage Employee Records		HR Department	HR department manages employee records

3.8. Customer service process

3.8.1. BPMN

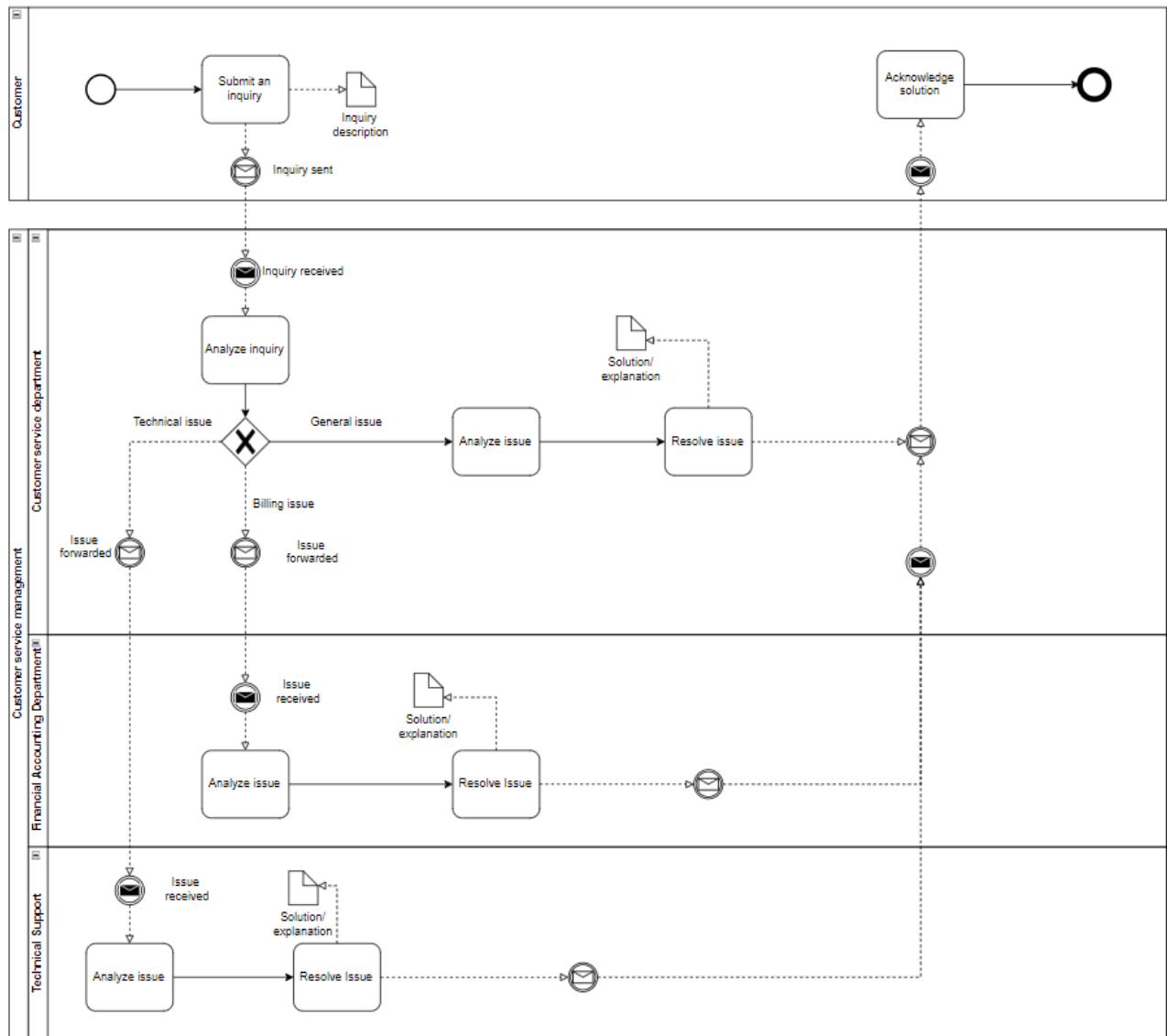


Figure 12: BPMN - Customer service process

3.8.2. Specification

Table 10: Customer service process specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Submit Inquiry		Customer	Customer submits an inquiry
Inquiry Received		Customer Service Department	The inquiry is received by the customer service department

Analyze Inquiry		Customer Service Department	The inquiry is analyzed to determine the nature of the issue
Technical Issue	Exclusive Gateway	Customer Service Department	The inquiry is related to a technical issue
General Issue	Exclusive Gateway	Customer Service Department	The inquiry is related to a general issue
Billing Issue	Exclusive Gateway	Customer Service Department	The inquiry is related to a billing issue
Forward to Technical Support		Customer Service Department	The inquiry is forwarded to the technical support team
Forward to Customer Service Management		Customer Service Department	The inquiry is forwarded to customer service management
Forward to Financial Accounting Department		Customer Service Department	The inquiry is forwarded to the financial accounting department
Issue Received		Technical Support / Customer Service Management / Financial Accounting Department	The relevant department receives the inquiry
Analyze Issue		Technical Support / Customer Service Management / Financial Accounting Department	The relevant department analyzes the issue
Resolve Issue		Technical Support / Customer Service Management / Financial Accounting Department	The relevant department resolves the issue

Provide Solution/Explanation		Technical Support / Customer Service Management / Financial Accounting Department	The relevant department provides a solution and explanation to the customer service department
Send Solution/Explanation to Customer		Customer Service Department	The customer service department sends the solution and explanation to the customer

3.9. AI implementation

3.9.1. BPMN

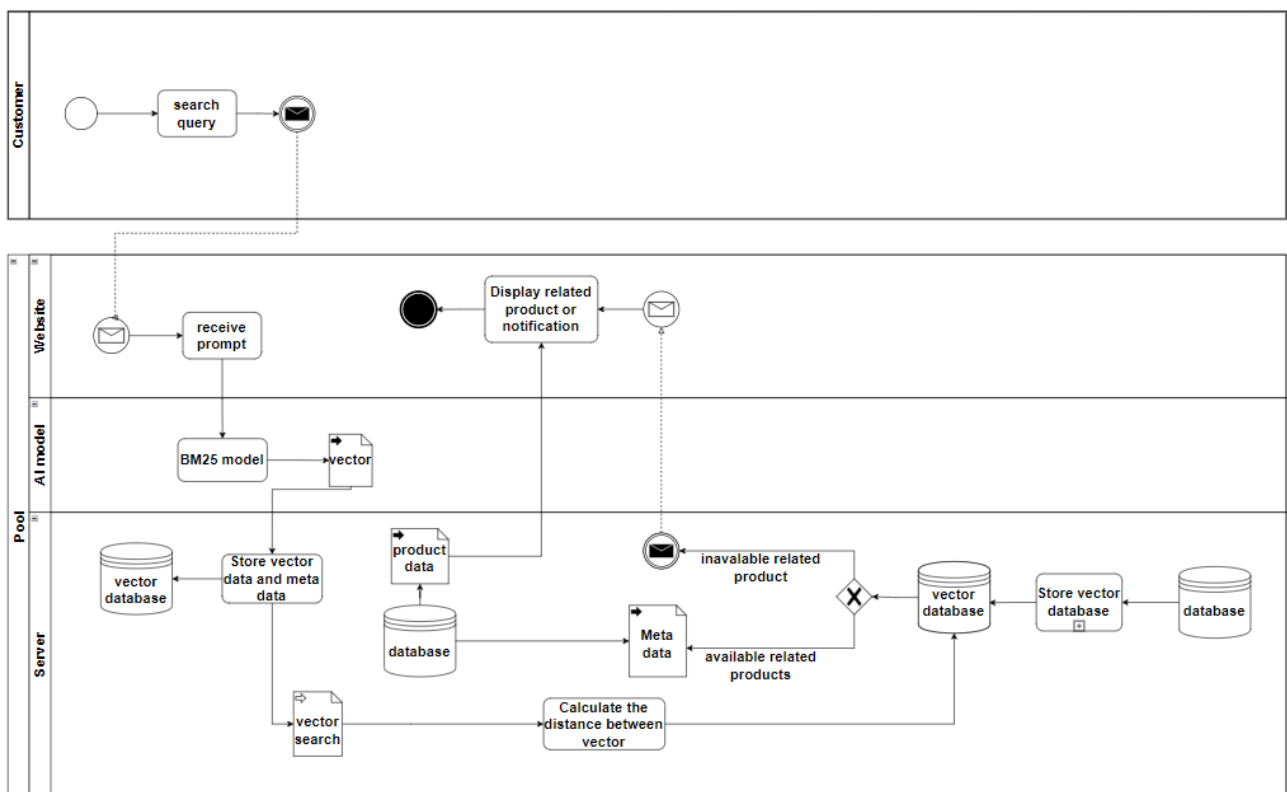


Figure 13: BPMN - AI implementation

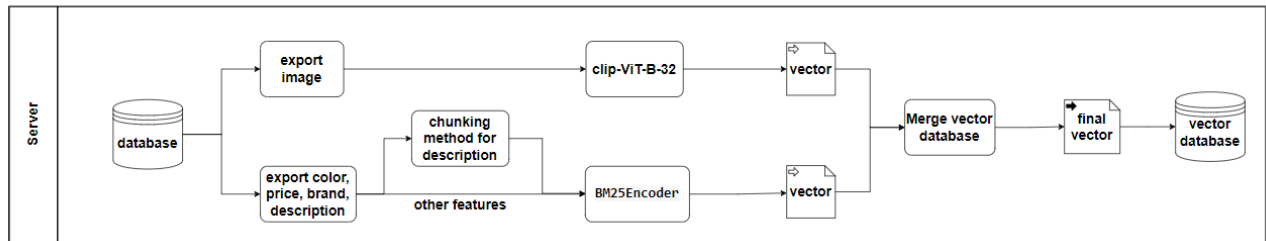
Store vector database

Figure 14: Store vector database process

3.9.2. Specification

Table 11: AI implementation specification

ACTIVITY	GATEWAY	PERFORMER	DESCRIPTION
Apply (for something)		Candidate	Candidate submits an application
Receive Application			Department/Role receives application
Conduct Pre-screening			Maybe HR or a designated team conducts a pre-screening process
Pre-screening Interview			Maybe an interview is conducted after pre-screening
Qualified (Yes)	Exclusive Gateway		Candidate moves forward in the process
Not Qualified (No)	Exclusive Gateway		Candidate doesn't move forward in the process
Send Rejection Letter (1st)			A rejection letter is sent to the candidate
Conduct Evaluation			Maybe a hiring manager or a team evaluates the candidate
Schedule Interview			An interview is scheduled if the candidate is chosen
Interview			The candidate is interviewed

Pass Interview	Exclusive Gateway		Candidate passes the interview
Fail Interview	Exclusive Gateway		Candidate fails the interview
Send Rejection Letter (2nd)			A rejection letter is sent to the candidate
Offer Job			The hiring manager or a designated role offers the job
Make Decision		Candidate	The candidate decides to accept or reject the job offer
Accept Job Offer	Exclusive Gateway		Candidate accepts the job offer
Decline Job Offer	Exclusive Gateway		Candidate rejects the job offer
Send Offer Letter			An offer letter is sent to the candidate
Receive Signed Offer			The department/role receives the signed offer letter
Onboarding			The HR department or a designated team completes the onboarding process

CHAPTER 4: PROPOSED ALGORITHMS AND SOLUTIONS

In Chapter 4, we examine the implementation of the CLIP model for semantic search within ERP systems, focusing on its integration into the sales process. This advanced machine learning algorithm, developed by OpenAI, processes both images and text, significantly enhancing search functionalities across diverse data formats. We outline the fundamental concepts of the CLIP model, detail the integration steps into ERP systems, and discuss the required architectural adjustments.

The chapter also highlights the benefits of semantic search, such as improved efficiency in customer service, enhanced product discoverability, and better user experience. Through practical examples and case studies, we illustrate how these improvements lead to better sales outcomes and greater customer satisfaction. By the end, readers will understand the strategic advantages of incorporating advanced semantic search technologies in optimizing ERP systems.

4.1. Introduction to dataset

4.1.1 Data collection

Data is an immensely crucial aspect in acquiring insights into a given issue, study, research, or even persons. This is why it is viewed as an essential component of all the systems that comprise our modern world.

Using Selenium to crawl data from a website involves automating a web browser to dynamically extract key product information, including product names, brands, prices, descriptions, and image links. This process begins with installing Selenium and the appropriate web driver, followed by scripting the browser to navigate to the target website. Selenium identifies and selects the relevant HTML elements through methods like XPath or CSS selectors. It then extracts the desired data fields and stores them in structured formats, in this case we store it in CSV format. Selenium's ability to interact with JavaScript-rendered content and handle complex web elements like dropdowns and infinite scrolls ensures accurate and comprehensive data collection, making it a powerful tool for obtaining up-to-date product information directly from live websites.

Due to the high demand for collecting product data, particularly information on clothing from various brands, we chose Maison.com – a widely recognized website that distributes numerous popular brands. Our focus is on gathering data for brands like FILA,

PUMA, MLB, and DSQUARED2. For each brand, we collect detailed information including the brand name, product name, price, color, and description. This structured approach allows us to create comprehensive data entries for each product, ensuring that our recommendation system can accurately reflect the diverse offerings of these brands. By leveraging Maison.com, we can efficiently gather high-quality data essential for our e-commerce platform's needs.

- **Brand:** The brand of the product.
- **Product:** Name of the product.
- **Price:** Price of the product.
- **Description:** A brief description of the product.
- **Color:** The primary color of the product.
- **Product Image:** The images of the product

ID	Brand	Product	Price	Description	Color
d59980fc-f88f-42c6-8122f-82	FILA	FILA Mindblower	499000	Embrace a blend of Black and Red	Black
2c68122f-82b1-4b18-53680-9e	FILA	FILA Skyjacker	350000	Transition from street to court	Black
b1853680-9e97-4e72-30c-50	FILA	FILA Disruptor	750000	Make a statement in the court	Red
97e7230c-50b6-4f1c-0f0-b8	FILA	FILA Renno	699000	Classic style meets modern	Blue
b6f1c0f0-b8f5-4569-51b4-bc	FILA	FILA Court Deluxe	720000	Own the court in the Fall	White
aacb3480-bf8f-48fa-88bc6-c6	MLB	MLB Basic Cotton	769000	Suitable for summer and fall	Beige
8fa88bc6-c6b5-4587-81f5-15	MLB	MLB Basic Cotton	769000	Dynamic, comfortable and stylish	Black
b58781f5-1551-4a90-29e-b4	MLB	MLB Basic Mega	990000	Comfortable and stylish	Black
51a9029e-b406-47e9-fb7-1f	MLB	MLB Basic Mega	799000	Comfortable and stylish	Melange Grey
067e9fb7-1fb6-4f1c-0f0-b8	MLB	MLB Checkerboard	799000	Trendy, overfit t-shirt	White
b6f1c0f0-b8f5-4569-51b4-bc	PUMA	PUMA Better Classics	499000	Classic round-neck t-shirt	Pink
f56951b4-bc4d-4f86-633a-8b	PUMA	PUMA Better Classics	499000	Durable and comfortable	White
4df8633a-8baacb-3480-bf	PUMA	PUMA Classics	499000	Simple and stylish t-shirt	Black
aacb3480-bf8f-48fa-88bc6-c6	PUMA	PUMA Classics	599000	Simple and stylish t-shirt	White
8fa88bc6-c6b5-4587-81f5-15	PUMA	PUMA Cloudspun	599000	High-performance running	Pink

Figure 15: Data set

So, what is the reason why we choose these features in the recommender system?

Brand: Brand loyalty significantly influences customer purchasing decisions. Including the brand helps the recommendation system suggest products from brands that customers prefer or have previously bought from, increasing the likelihood of a successful recommendation.

Product: The product name or model contains keywords that can indicate the type of product and its category. This helps in identifying and recommending similar or complementary products, enhancing the relevance of the recommendations.

Price: Price is a critical factor in customer decisions. By considering price, the recommendation system can suggest products within a similar budget range to those the customer has shown interest in, ensuring the recommendations are both relevant and affordable.

Description: Descriptions provide detailed insights into the product's features, benefits, and use cases. Vectorize techniques can analyze these descriptions to find similarities between products, improving the accuracy and relevance of recommendations.

Color: Aesthetic preferences, including color, play a significant role in purchasing decisions, especially in the fashion industry. Including color allows the system to account for customer preferences and suggest products in colors they are likely to prefer, enhancing customer satisfaction.

By incorporating these features, our recommendation system can deliver personalized and relevant product suggestions, ultimately improving the customer experience and increasing engagement on the platform.

Data Preprocessing

Handling Missing Values: we started by identifying any missing values in the dataset. This includes checking for null entries in critical fields such as brand, product name, price, description, and color. For non-critical fields with occasional missing values, we used imputation techniques. For instance, missing colors might be inferred from similar products. However, if key fields like product name or brand were missing, the record was excluded to maintain data integrity.

```
[ ] df = pd.DataFrame(data)
    # Delete rows with any missing values
    df.dropna(inplace=True)
```

Figure 16: Delete data line that has missing values

Normalizing Text Fields: All text fields were converted to lowercase to ensure consistency in comparisons and searches (e.g., "BLACK" and "black" are treated the same). Unnecessary special characters and punctuation were removed from product names and descriptions to simplify text processing. Extra spaces and line breaks were removed to ensure clean and readable text fields.

```
dataset['Description'] = dataset['Description'].str.lower()
dataset['Color'] = dataset['Color'].str.lower()
dataset['Brandname'] = dataset['Description'].str.lower()
dataset['ProductName'] = dataset['Description'].str.lower()
dataset.head()
```

Figure 17: Convert all text to lower characters

```
# Function to remove special characters and punctuation
def clean_text(text):
    if text:
        # Remove special characters and punctuation
        text = re.sub(r'^\w\s', '', text)
        return text.strip()
    else:
        return None
```

Figure 18: Delete the special characters

By combining product information, we gain valuable insights about products. This integrated comprehensive details such as brand, product name, price, color, and description. This combination allows us to analyze patterns and trends, enabling the hybrid search system to make more accurate and personalized requirements. By understanding what customers expect to see, we can predict which products they might be interested in, thereby enhancing the overall shopping experience and increasing customer satisfaction.

4.2. Introduces the application of deep learning to the process

In the realm of semantic search within ERP systems, the utilization of the CLIP model holds promise for optimizing the sales process while addressing challenges. This model offers a multifaceted approach by enhancing customer experience, streamlining resource allocation, and expediting system responsiveness. Moreover, it effectively

addresses issues such as irrelevant data and data distortion encountered during sales procedures.

The benefits of the CLIP model in semantic search for customers stem from its multifaceted approach. By leveraging advanced algorithms and deep learning techniques, the model analyzes vast amounts of data to understand customer preferences and behavior. This enables it to offer personalized recommendations tailored to individual needs, thereby enhancing the overall customer experience. [1]

Additionally, by optimizing resource allocation, the model ensures that the right resources are allocated to the right tasks at the right time, minimizing inefficiencies and maximizing productivity. Moreover, its ability to expedite system responsiveness is rooted in its capacity to quickly process and retrieve relevant information, leading to faster and more accurate responses to customer queries. Overall, these capabilities translate into tangible benefits for customers, including improved satisfaction, efficiency, and effectiveness in their interactions with the ERP system. [2]

The problem of irrelevant data can be handled very smoothly because of the combination of all the text data that describes the product and the specific characteristics. This means that the model does not trust only the text description or the image of product, but the result must be the solution of both text and image data. [3]

Conclusion, the CLIP model can contribute a lot to ERP system especially sale process which provides many advantages for both customers and businesses. This is also the method that many E-commerce platforms have been developing.

4.3. Implementation steps

Hybrid semantic search is a complex process that requires careful planning and implementation. To simplify this project, we divided it into 4 steps:

1. Combine retrieval using clip-ViT-B-32 model and BM25 Retriever: This step involves using two different data retrieval methods, including clip-ViT-B-32 and BM25 Retriever, to find information that matches the user's request.
2. Process the output vector and store in vector database: This step takes responsibility for storing sparse vector and dense vector in vector database.

3. Processing data after retrieval: In this step, data after being retrieved will be processed through K-nearest neighbor algorithm to optimize the results returned to the user.
4. Optimize search engine: we must use the method called merging sparse vector and dense vector using hyper parameter.

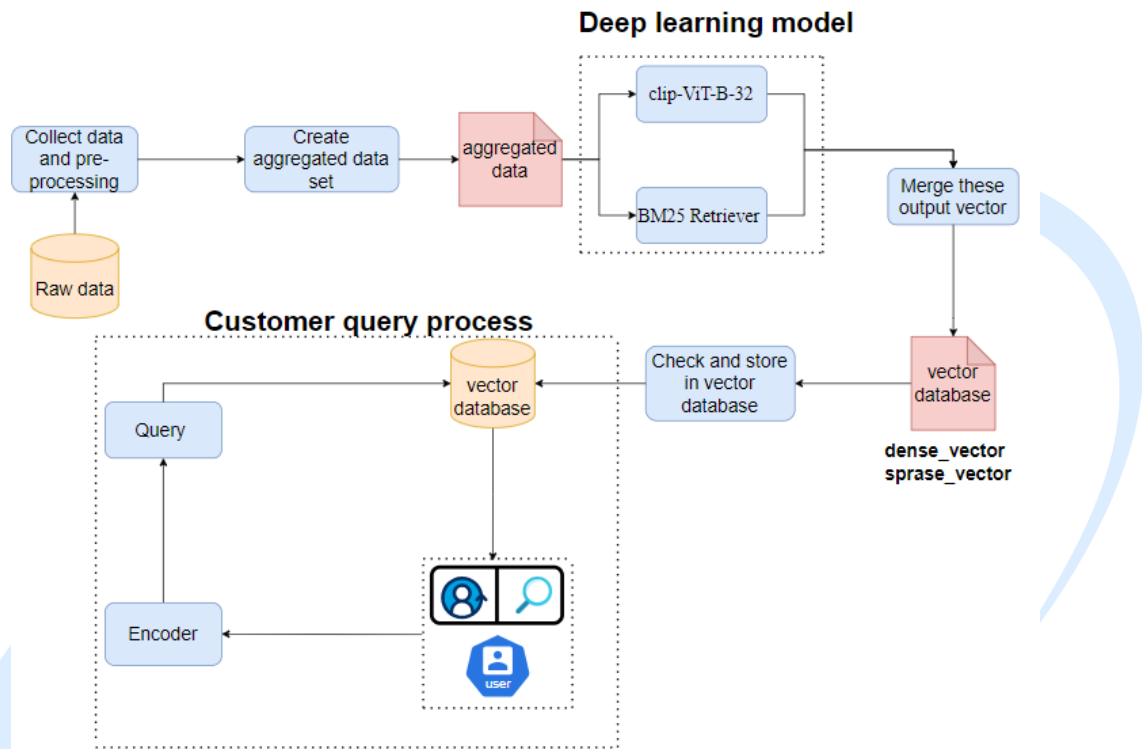


Figure 19: Architecture of AI model and how the user interact with customer

4.3.1. Combine retrieval using clip-ViT-B-32 model and BM25 Retriever

In this stage, we employ two distinct models to transform the collected data into vectors. For image data, we utilize the CLIP-ViT-B-32 model, while for textual information—such as descriptions, colors, brand names, product names, and prices—we use the BM25 algorithm. [4]

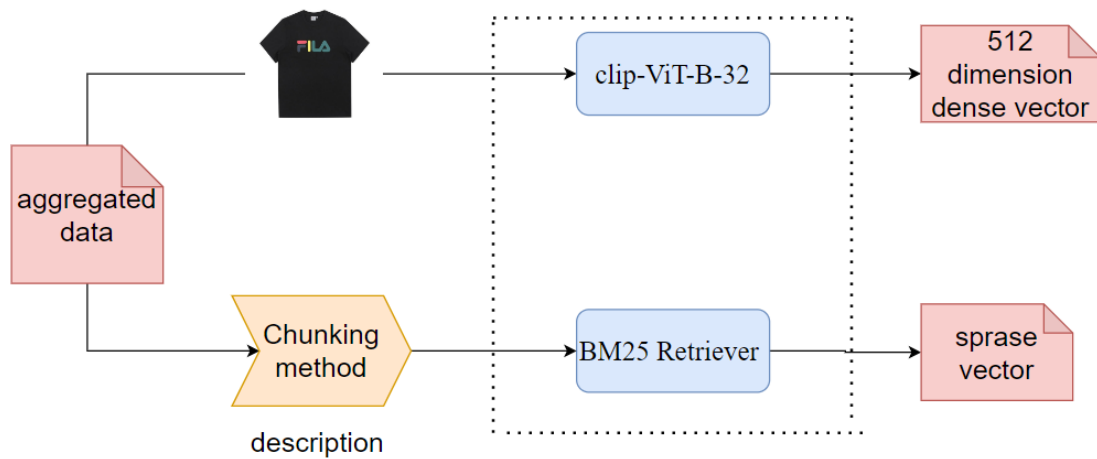


Figure 20: Convert from raw data to vectors

In many cases, the product description may be too lengthy to convert into a single vector. Therefore, we decided to split the description into smaller segments and convert each segment into individual vectors called chunking method. Afterward, these vectors are combined to form a comprehensive description vector for each product. The final dense vector for each product is created by merging the vectors for the color, brand, and product name. This final vector is then stored as a sparse vector in our database.

```
[ ] from pinecone_text.sparse import BM25Encoder

bm25 = BM25Encoder()
bm25.fit(metadata['productDisplayName'])
```

100% 440

Figure 21: Fit BM25 model with our data

```
{'indices': [23789636,
1830646559,
632192512,
931643408,
3905155331,
3828986392],
'values': [0.4449638258432887,
0.4449638258432887,
0.4449638258432887,
0.4449638258432887,
0.4449638258432887,
0.4449638258432887]}
```

Figure 22: Sparse vector of each data point

For images, we use the CLIP-ViT-B-32 model to create dense vectors that encapsulate product features the model can analyze, such as patterns, motifs, and designs, etc

```

▶ from sentence_transformers import SentenceTransformer
import torch

device = 'cuda' if torch.cuda.is_available() else 'cpu'

# load a CLIP model from huggingface
model = SentenceTransformer(
    'sentence-transformers/clip-ViT-B-32',
    device=device
)
model

```

⇌ SentenceTransformer(
 (0): CLIPModel()
)

Figure 23: CLIP-ViT-B-32 - convert the image to dense vector

4.3.2. Process the output vector and store in vector database

We store both dense and sparse vectors in the Pinecone database, a robust vector database that allows us to query using various metrics such as cosine similarity, Manhattan distance, and more. [5]

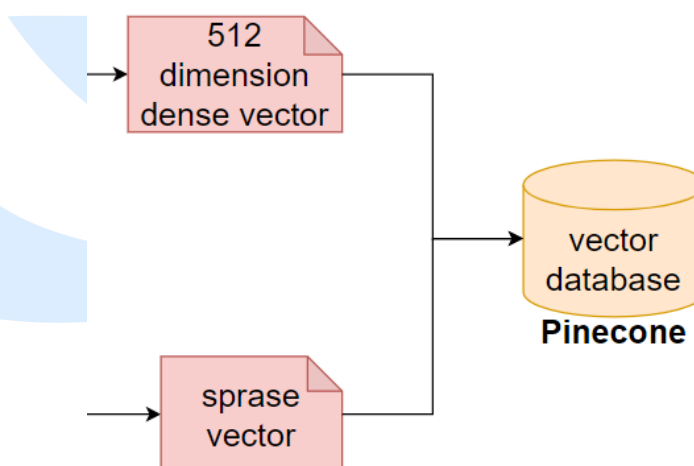


Figure 24:Store data in vector database Pinecone

The Pinecone database also allows us to store meta data for each product.

ID	VALUES
33	-0.0351644605, 0.0933295861, -0.0266190823, 0.0276485607, 0.0328844339, -0.0127358...

METADATA

Color: "White"

Description: "color: deep red, white, dark navy\nmaterial: 100% cotton\ndynamic loose-fitting t-shirt design\nround neck"

Price: "895000"

Product Name: "Heritage Loose Fit Logo round neck short sleeve unisex t-shirt T1"

sku_id: "c97523c2-d7aa-4007-957b-783503eb76aa"

Figure 25: Data that stored in Pinecone including vector and metadata

4.3.3. Processing data after retrieval

When a customer sends a request to our application, the query is encoded into a vector that captures the customer's demands and expectations.

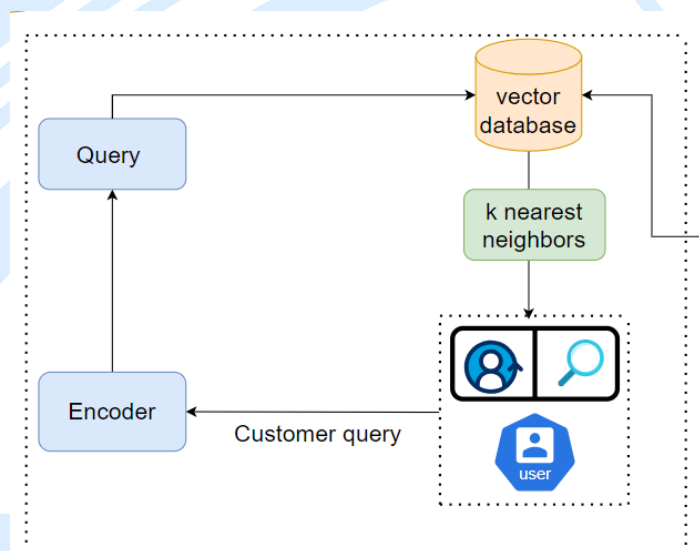


Figure 26: Architecture that customer interact with system

In our case, the vector encoded by BM25 from the customer's prompt will be compared with the vectors stored in the database using a distance algorithm, such as the dot product. The function will then output the top k most relevant products, along with their metadata and IDs.

```

query = "dark blue french connection jeans for men"

# create sparse and dense vectors
sparse = bm25.encode_queries(query)
dense = model.encode(query).tolist()
# search
result = index.query(
    top_k=14,
    vector=dense,
    sparse_vector=sparse,
    include_metadata=True
)

```

```

French Connection Men Blue Jeans
French Connection Men Blue Jeans
French Connection Men Blue Jeans
French Connection Men Blue Jeans
French Connection Men Blue Jeans
French Connection Men Blue Jeans
French Connection Women Blue Jeans
French Connection Women Blue Jeans
French Connection Men Navy Blue Jeans
French Connection Men Blue paint Stained Regular Fit Jeans
French Connection Men Black Jeans
French Connection Men Black Jeans
French Connection Men Black Jeans
French Connection Men Black Jeans

```



4.3.4. Optimize search engine

In the above function, we implemented a simple query to calculate the distance between vectors. The results are generally acceptable, but in some cases, such as when searching for "dark blue French connection jeans for men," the function returns jeans but fails to correctly rank men's French Connection jeans, and some women's jeans are included.

This issue arises because the model primarily considers the image of each product. To address this, we decided to build a function that prioritizes the balance between dense and sparse vectors in our search engine. This function, named **hybrid_scale**, allows us to adjust the priority between these vectors using a hyperparameter called alpha, which defines the extent to which the image data (image vector) will be utilized: [6]

```

def hybrid_scale(dense, sparse, alpha: float):
    if alpha < 0 or alpha > 1:
        raise ValueError("Alpha must be between 0 and 1")
    # scale sparse and dense vectors to create hybrid search vecs
    hsparse = {
        'indices': sparse['indices'],
        'values': [v * (1 - alpha) for v in sparse['values']]
    }
    hdense = [v * alpha for v in dense]
    return hdense, hsparse

```

Test for alpha = 0:

```
question = "dark blue french connection jeans for men"

# scale sparse and dense vectors
hdense, hsparse = hybrid_scale(dense, sparse, alpha=0)
# search
result = index.query(
    top_k=14,
    vector=hdense,
    sparse_vector=hsparse,
    include_metadata=True
)
# used returned product ids to get images
imgs = [images[int(r["id"])] for r in result["matches"]]
# display the images
display_result(imgs)
```

French Connection Men Blue Jeans
 French Connection Men Blue Jeans
 French Connection Men Blue Jeans
 French Connection Men Blue Jeans
 French Connection Men Blue Jeans
 French Connection Men Blue Jeans
 French Connection Women Blue Jeans
 French Connection Women Blue Jeans
 French Connection Men Navy Blue Jeans
 French Connection Men Blue paint Stained Regular Fit Jeans
 French Connection Men Black Jeans
 French Connection Men Black Jeans
 French Connection Men Black Jeans
 French Connection Men Black Jeans



The same result as above because alpha value equal to zero which means there is nothing changed compared to the above result.

Test for alpha = 1:

```
# scale sparse and dense vectors
hdense, hsparse = hybrid_scale(dense, sparse, alpha=1)
# search
result = index.query(
    top_k=14,
    vector=hdense,
    sparse_vector=hsparse,
    include_metadata=True
)
# used returned product ids to get images
imgs = [images[int(r["id"])] for r in result["matches"]]
# display the images
display_result(imgs)
```

Locomotive Men Radley Blue Jeans
 Locomotive Men Race Blue Jeans
 Locomotive Men Eero Blue Jeans
 Locomotive Men Cam Blue Jeans
 Locomotive Men Ian Blue Jeans
 French Connection Men Blue Jeans
 Locomotive Men Cael Blue Jeans
 Locomotive Men Lio Blue Jeans
 French Connection Men Blue Jeans
 Locomotive Men Rafe Blue Jeans
 Locomotive Men Barney Grey Jeans



Because an alpha value equal to 1 means the function only considers image data (image vector), the brand names of the results may be incorrect. Therefore, we need to test

multiple times to find a reasonable alpha value that balances the performance of both image and text vectors.

Test for alpha = 0.05

```
# scale sparse and dense vectors
hdense, hparsed = hybrid_scale(dense, sparse, alpha=0.05)
# search
result = index.query(
    top_k=14,
    vector=hdense,
    sparse_vector=hparsed,
    include_metadata=True
)
# used returned product ids to get images
imgs = [images[int(r["id"])] for r in result["matches"]]
# display the images
display_result(imgs)
```

```
French Connection Men Blue Jeans
French Connection Men Blue Jeans
French Connection Men Blue Jeans
Locomotive Men Radley Blue Jeans
French Connection Men Navy Blue Jeans
Locomotive Men Race Blue Jeans
Locomotive Men Eero Blue Jeans
Locomotive Men Cam Blue Jeans
French Connection Men Blue Jeans
French Connection Men Blue paint Stained Regular
```

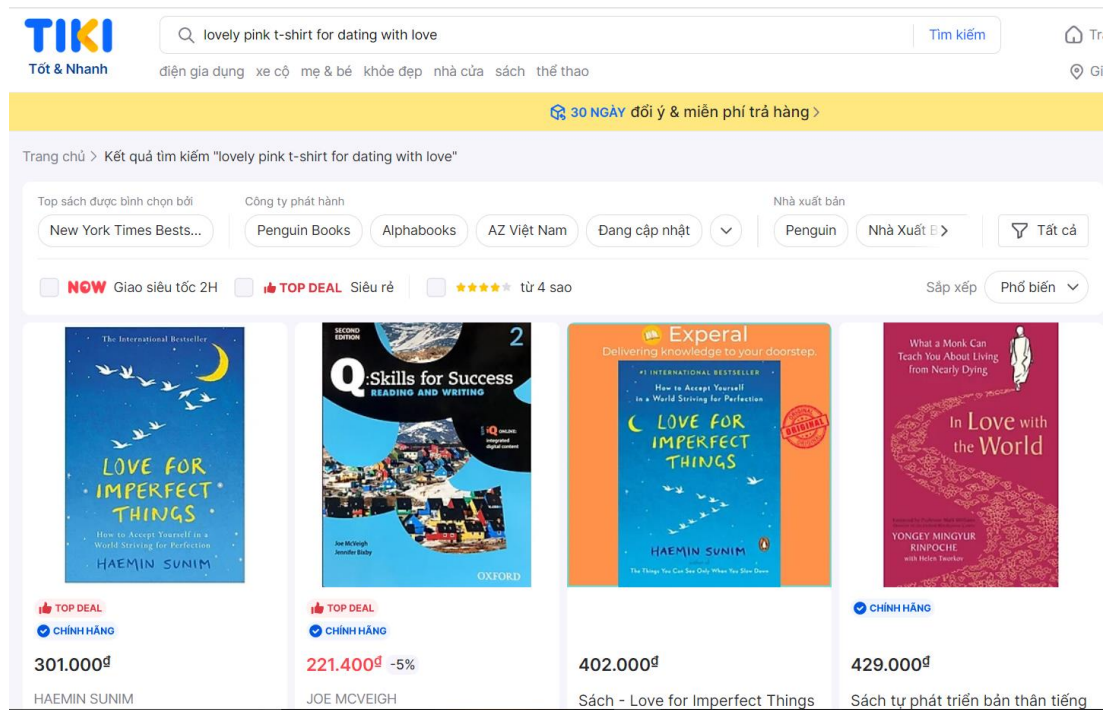


With this approach, the model predominantly recommends men's jeans from French Connection but also includes products from other brands. This not only ensures that the customer receives the desired product but also offers additional related products, providing the customer with more options. [6]

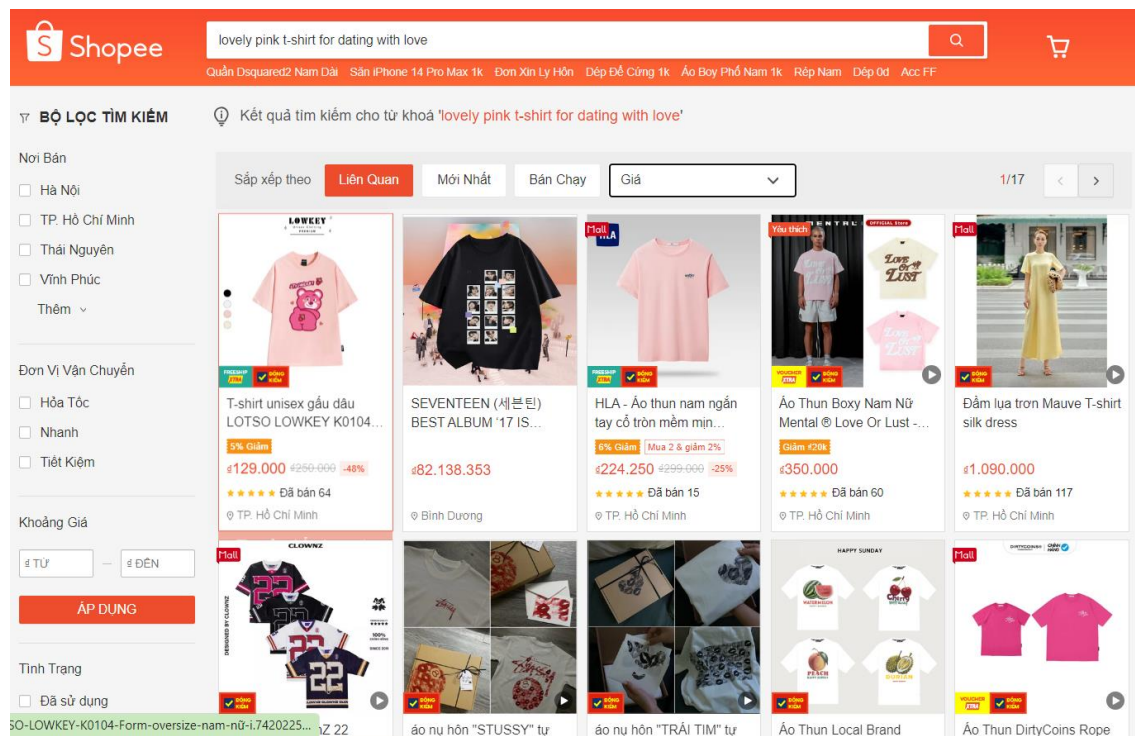
In conclusion, to enhance the search results, we can adjust the hyperparameter, namely the alpha value, to prioritize either the text vector or the image vector. By doing so, we can better meet the expectations of customers in various scenarios. This adjustment ensures that the search engine delivers more accurate and relevant results, thereby enhancing the overall customer experience.

4.3.5. Experiment on our e-commerce website

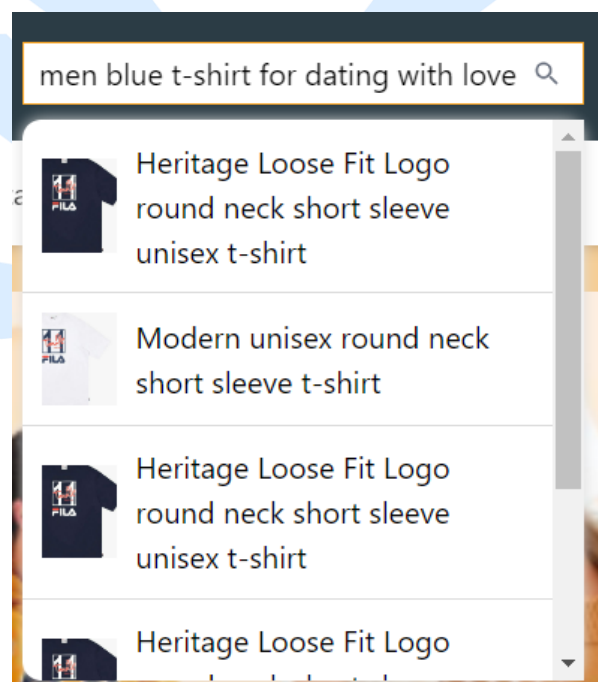
In this case, Tiki still struggles with addressing the semantic search needs of its customers. For example, when a user searches for "*lovely pink t-shirt for dating with love*" Tiki fails to understand the specific request and returns irrelevant products that do not match the customer's intent. This issue highlights the limitations in Tiki's search functionality, where the platform does not accurately interpret and respond to more nuanced and detailed search queries.



In this case, a similar issue occurs with Shopee. When the same query, "*men blue t-shirt for dating with love*," is entered, Shopee does return products that are indeed shirts. However, its search tool fails to recognize the specific requirements regarding color and the context of the query. This demonstrates a limitation in Shopee's search functionality, where it cannot fully interpret and meet the nuanced needs of the customer's search intent.



A search tool leveraging the CLIP model, which combines information from both images and product descriptions, can address this issue effectively. This model can understand the semantics of the customer's request and identify key elements within the query. By doing so, it can provide products that accurately meet the customer's needs, something that leading e-commerce platforms in Vietnam have yet to achieve.



CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

This chapter gives a summary of the main findings of the research based on the result in chapter 4. It hi

ghlights the meaning of our project to the business and practical applications as well as suggests missions for the future development based on the objectives that is not completed.

5.1. Advantages

The implementation of a recommendation system and semantic search offers numerous advantages for businesses, particularly for the International Brand Distributor Store System.

Recommendation systems make shopping **more enjoyable for customers**. These systems use smart algorithms to figure out what customers like, based on things like what they've looked at before and what they've bought. This way, the system can recommend products that are a good fit for each customer. This not only saves customers time searching for things they might like, but it also makes them happier with the shopping platform overall and more likely to keep coming back. [7]

Besides, personalized product suggestions **increase the likelihood of purchases**. When customers are presented with products that match their preferences and needs, they are more likely to buy. This capability is especially potent in e-commerce, where customer retention and repeated purchases are crucial for business growth.

Built to handle more customers and sales, the recommendation system can grow alongside the business. This flexibility guarantees the platform stays **strong and dependable**, even during busy periods, ultimately contributing to the business's success in the long run. The implementation of semantic search using the CLIP model significantly enhances data retrieval processes. Unlike traditional keyword-based searches, semantic search understands the context and intent behind queries, providing more accurate and relevant results. This leads to a more intuitive and efficient search experience for users. [8]

Semantic search allows users to employ natural language queries, which are matched semantically with relevant data. This means that users can describe what they are looking for in everyday language, and the system will understand and retrieve the most relevant information, whether it is text documents or images. This user-friendly approach **reduces frustration and encourages** more frequent use of the ERP system. By enabling more intuitive search capabilities, semantic search helps users find the most relevant and accurate information quickly. This supports better decision-making, as users can rely on comprehensive and precise data. [9]

5.2. Limitations

While recommendation systems and semantic search have clear advantages, research and implementation revealed some drawbacks that need to be addressed.

For recommendation systems and semantic search to work well, they **need a lot of good data**. If the data is bad or there isn't enough of it, the recommendations and search results will be wrong, which can hurt customer satisfaction and sales. These systems are also **expensive and complex** to build and keep running, which can be a problem for small businesses.

Using customer data raises privacy concerns, so companies need to be careful about following **data protection** laws and keeping their **customers' trust**. Also, the algorithms can be biased, favoring certain things over others. Companies need to keep an eye on this and adjust to avoid unfair recommendations. Finally, integrating semantic search with other business systems can be technically difficult. It takes a lot of work and expertise to make sure everything works together smoothly.

5.3. Future directions

Because of the numerous benefits given, semantic search, a hybrid search that uses both images and text, is in high demand at the point of sale in many ERP systems today. Don't only use the search system; our team opted to present the metadata to the management system and integrate the recommendation system to improve the customer experience.

The main ideas of our future are building the end-to-end production project integrating semantic search, recommendation system and Llama2. This is very strong functions which supports a lot for managing system as well as system.

Llama 2 is an advanced chatbot designed to streamline managerial and financial inquiries within an organization. By leveraging real-time data analytics, Llama 2 provides managers with instant, comprehensible responses to complex questions about systems and finances. This capability eliminates the need for direct interaction with backend systems, saving both time and resources. Additionally, on the customer side, Llama 2 enhances the shopping experience by efficiently guiding users to their desired products. This not only reduces the time customers spend searching but also increases their likelihood of making additional purchases. Moreover, the reliability and user-friendly interface of Llama 2 foster a sense of trust and satisfaction among users. [10]

We aim to make a system which incorporates an information retrieval mechanism that sources relevant data from both internal and external knowledge bases, such as Wikipedia or industry-specific databases. Upon receiving a query, this system initially searches for pertinent documents or text segments in the database that align with the question's content. The retrieved documents are then amalgamated with the original query to establish a comprehensive context for the Large Language Model (LLM). This enriched context, which includes both the initial inquiry and the information gleaned from external sources, is processed by the LLM to produce the final response. Consequently, this integration allows the LLM to utilize a blend of internal knowledge and external data, enhancing the factual accuracy and user-friendliness of its responses. [11]

In conclusion, the integration of semantic search, the recommendation system, and Llama 2 into our ERP systems signifies a transformative step towards enhancing both managerial efficiency and customer satisfaction. By combining advanced search capabilities with dynamic data analytics and user-friendly interfaces, this comprehensive system not only streamlines internal processes but also significantly improves the customer journey. Future enhancements will focus on refining these integrations, expanding our database sources, and improving real-time response capabilities, thereby ensuring that our system remains at the forefront of technological innovation. Through continuous development and integration of these sophisticated tools, we aim to set new standards in ERP system functionality, making our platform a benchmark for industry excellence.



LIST OF REFERENCES

- [1] Yu, W., Liu, Y., Hua, W., Jiang, D., Ren, B., & Bai, "Turning a clip model into a scene text detector," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.*, pp. 6978-6988, 2023.
- [2] Garg, S., Farajtabar, M., Pouransari, H., Vemulapalli, R., Mehta, S., Tuzel, "Tic-clip: Continual training of clip models," 2023.
- [3] Sun, Q., Fang, Y., Wu, L., Wang, X., & Cao, Y, "Eva-clip: Improved training techniques for clip at scale," 2023.
- [4] Ao, Tenglong, Zeyi Zhang, and Libin Liu., "Gesturediffuclip: Gesture diffusion model with clip latents," pp. *ACM Transactions on Graphics (TOG)*, 42(4), 1-18., 2023.
- [5] Falcon, Alex, Giuseppe Serra, and Oswald Lanz, " "Improving semantic video retrieval models by training with a relevance-aware online mining strategy."," p. *Computer Vision and Image Understanding* (2024): 104035..
- [6] Zhao, Yida, Yuqing Song, and Qin Jin., ""Progressive learning for image retrieval with hybrid-modality queries."," p. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval.*, 2022.
- [7] Taipalus, Toni., ""Vector database management systems: Fundamental concepts, use-cases, and current challenges."," p. *Cognitive Systems Research* (2024): 101216..
- [8] VM, Kushala, Harikrishna Warriar, and Yogesh Gupta., ""Fine Tuning LLM for Enterprise: Practical Guidelines and Recommendations."," p. *arXiv preprint arXiv:2404.10779* (2024)..
- [9] Engels, J., Landrum, B., Yu, S., Dhulipala, L., & Shun, J., "Approximate Nearest Neighbor Search with Window Filters.," p. *arXiv preprint arXiv:2402.00943.*, 2024.
- [10] "Gao, Y., Arava, S. K., Li, Y., & Snyder Jr, J. W," *Improving the Capabilities of Large Language Model Based Marketing Analytics Copilots With Semantic Search And Fine-Tuning.*, p. *arXiv preprint arXiv:2404.13077.*, 2024.
- [11] MOSKVORETSKII, Viktor, et al. TaxoLLaMA, "WordNet-based Model for Solving Multiple Lexical Sematic Tasks," pp. *arXiv preprint arXiv:2403.09207.*, 2024.