GRADUATE CERTIFICATE IN INTELLIGENT REASONING SYSTEM (IRS)

Project Report

Know Your Coffee

Group Members:

Jingyi Zhang (A0269366M)

Fong Zhi En Kelvin (A0269365N)

Huiwen Tang (A0269361X)

Gengwei Li (A0269362W)

Lin Ji (A0269369H)

Table of Contents

1. Introduction

In these years, coffee has become one of the most important things to modern people, serving as both a social and functional staple in everyone's daily life. The emergence of coffee culture, the impact of globalization, and the recognized health advantages of moderate coffee consumption all contribute to its growing prominence. From a data-driven perspective, the growing importance of coffee in modern society can be observed through several indicators. These include the increasing number of coffee consumers, the rising consumption per capita, the growth of the coffee industry, and the expanding market for specialty coffee. As shown in Fig 1, according to the International Coffee Organization (ICO) [1], global coffee consumption has been steadily increasing over the years, demonstrating the expanding role of coffee in the daily lives of people worldwide. Also as shown in Fig 2 [2, 3], from ICO's data indicates that the per capita consumption of coffee has been on the rise for the past 30 years. This suggests that coffee is not only maintaining its importance among some traditional coffee countries but also gaining popularity among new coffee countries.

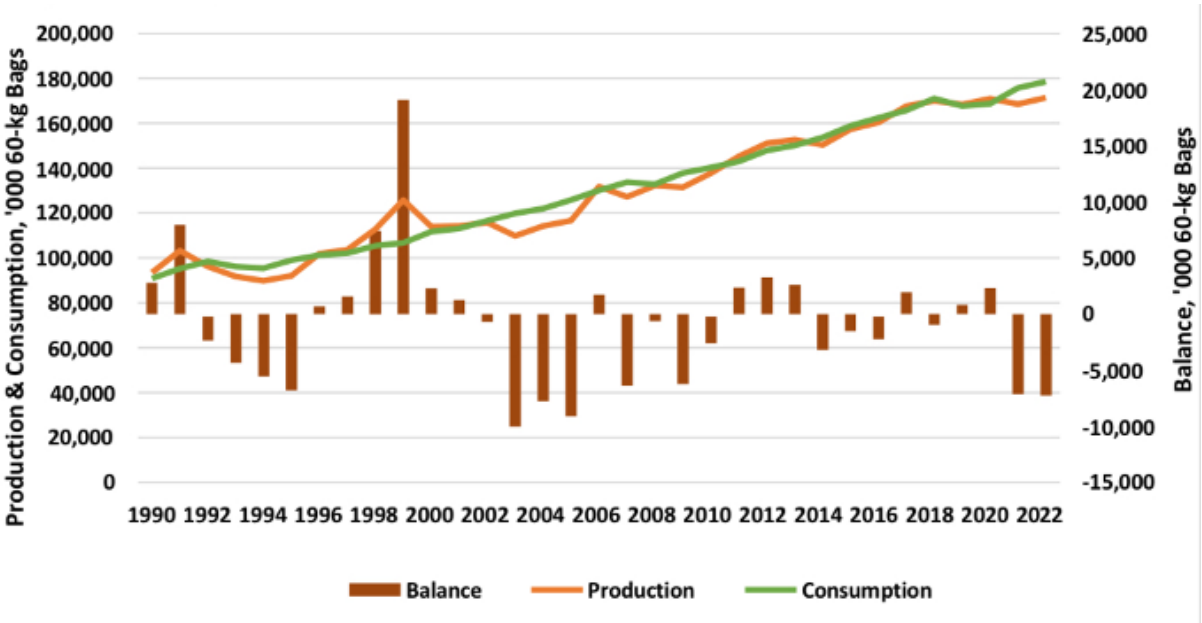

Figure 1. Summary of the World Coffee Market

World coffee consumption
In thousand 60kg bags

	2017/18	2018/19	2019/20	2020/21	CAGR (2017/18-2020/21)
World	**161 377**	**168 492**	**164 202**	**166 346**	**1.0%**
Africa	11 087	12 017	12 024	12 242	3.4%
Asia & Oceania	34 903	36 472	36 002	36 503	1.5%
Central America & Mexico	5 273	5 431	5 327	5 364	0.6%
Europe	53 251	55 637	53 372	54 065	0.5%
North America	29 941	31 779	30 580	30 993	1.2%
South America	26 922	27 156	26 898	27 180	0.3%
Exporting countries (Crop years)	***49 686***	***50 245***	***49 982***	***50 666***	***0.7%***
Brazil	21 997	22 200	22 000	22 400	0.6%
Indonesia	4 750	4 800	4 806	5 000	1.7%
Ethiopia	3 643	3 685	3 781	3 798	1.4%
Philippines	3 180	3 300	3 250	3 312	1.4%
Viet Nam	2 500	2 600	2 650	2 700	2.6%
Mexico	2 400	2 450	2 425	2 420	0.3%
Colombia	1 793	1 791	2 025	2 045	4.5%
Venezuela	1 600	1 550	1 275	1 100	-11.7%
India	1 470	1 475	1 450	1 485	0.3%
Thailand	1 375	1 400	1 400	1 415	1.0%
Guatemala	395	395	393	403	0.7%
Dominican Republic	390	390	377	383	-0.6%
Madagascar	365	375	375	377	1.1%
Honduras	375	375	350	375	0.0%
Costa Rica	353	365	352	353	0.0%
Haiti	340	345	345	341	0.1%
Côte d'Ivoire	317	317	317	317	0.0%
El Salvador	300	300	292	295	-0.6%
Peru	250	250	250	250	0.0%
Uganda	245	250	254	263	2.4%
Others	1 647	1 631	1 615	1 634	-0.3%
	111 691	***118 247***	***114 220***	***115 680***	***1.2%***
Importing countries (Coffee years: October - September)					
European Union	40 491	41 768	39 758	40 251	-0.2%
United States of America	26 112	27 759	26 651	26 982	1.1%
Japan	7 750	7 561	7 355	7 386	-1.6%
Russian Federation	4 324	4 691	4 631	4 681	2.7%
Canada	3 829	4 020	3 929	4 011	1.6%
Republic of Korea	2 371	2 476	2 471	2 513	2.0%
Australia	1 854	1 961	1 939	1 962	1.9%
Algeria	1 911	2 150	2 110	2 131	3.7%
Turkey	1 376	1 740	1 711	1 754	8.4%
Saudi Arabia	1 275	1 266	1 241	1 253	-0.6%
Ukraine	1 252	1 379	1 379	1 379	3.3%
Switzerland	1 013	1 079	1 060	1 074	2.0%
Norway	729	785	906	924	8.2%
Morocco	740	755	753	780	1.8%
Taiwan	690	703	707	725	1.6%
Lebanon	610	621	552	452	-9.5%
South Africa	638	673	653	655	0.9%
Egypt	614	1 209	1 242	1 279	27.7%
Argentina	604	623	621	644	2.2%
Sudan	605	741	713	702	5.1%
Others	12 904	14 288	13 838	14 144	3.1%

Data as at May 2021. Next update August 2021.

Figure 2. Table of World Coffee Consumption

Besides these, the growth of the coffee industry and expanding specialty coffee market are strong indicators for a healthy coffee growth. Overall, these data-driven insights underscore

the growing importance of coffee in contemporary society, as people continue to embrace it as a vital part of their daily routines, social lives and culinary adventures.

After studying that the world of coffee continues to grow in both variety and complexity, it becomes increasingly challenging for consumers to navigate the vast landscape of options and discover their personal preferences. Therefore, this coffee recommendation system idea came out. We want to design a coffee recommendation system that offers a solution to solve the 'Choosing best coffee' problem by leveraging advanced algorithms and data analysis to curate personalized coffee suggestions tailored to individual tastes and preferences. This innovative approach can not only enhance the coffee experience for the customers, but also drive engagement and customer loyalty for coffee retailers and subscription services.

In recent years, the preference for filter and espresso coffee over packaged coffee (eg. Instant coffee or canned coffee) has increased for several reasons. Firstly, filter and espresso coffee offer a superior taste and flavor experience compared to packaged coffee. Filter coffee, brewed using fresh ground beans and proper extraction techniques, allows for a more nuanced and well-balanced cup of coffee. The slower extraction process in filter coffee brings out the subtle flavors and aromas of the beans, resulting in a more enjoyable and satisfying drinking experience. Similarly, espresso coffee, with its concentrated and bold flavors, provides a rich and intense taste that many coffee enthusiasts appreciate. Secondly, filter and espresso coffee offer a greater degree of customization and personalization. Coffee shops often provide a wide range of specialty beans and brewing methods, allowing customers to select the specific type of coffee, roast level, and flavor profile that suits their preferences. This level of choice and control over the brewing process allows individuals to tailor their coffee to their liking, enhancing their enjoyment and satisfaction. Additionally, there has been a growing interest in the quality and origin of coffee beans. Many people are becoming more conscious of the sourcing practices, sustainability, and ethical considerations associated with coffee production. Filter and espresso coffee, particularly in specialty coffee shops, often prioritize high-quality beans that are ethically sourced and sustainably grown. This emphasis on quality and responsible sourcing resonates with consumers who value transparency and want to support environmentally and socially responsible practices.

Filter and espresso coffee align with this trend by highlighting the artistry and skill involved in preparing a great cup of coffee. This appeals to individuals who view coffee as more than just a caffeinated beverage but as a sensory experience and an opportunity to appreciate the craftsmanship of skilled baristas. Overall, the preference for filter and espresso coffee over packaged coffee has grown due to the superior taste, customization options, emphasis on quality and sustainability, and the evolving coffee culture that values craftsmanship and the coffee-drinking experience. So the next question will be how to choose suitable coffee in the Cafe, even if this is your first time visiting the shop.

A well-designed coffee recommendation system takes into account various factors such as the user's flavor preference, roast levels, brewing methods, filter or espresso (with or without milk), and even what time of the day they want to drink the coffee, to create a tailored list of coffee options. By employing techniques such as text similarity mining, collaborative filtering, content-based filtering, and hybrid approaches, the system can provide users with a diverse selection of coffee options that align with their tastes and preferences. On the other hand, in an increasingly competitive coffee market, the implementation of a coffee recommendation system not only benefits consumers by simplifying their decision-making process, but also offers a valuable tool for businesses to differentiate themselves and foster long-lasting relationships with their customers. As the demand for personalized experiences grows, a coffee recommendation system represents a key innovation in catering to the evolving needs of the modern coffee enthusiast.

2. Business Case

The Community Coffee located in Singapore [Fig 3], a company known for its commitment to sourcing, roasting, brewing, and serving high-quality coffees, is seeking to enhance its customer experience and market differentiation by implementing an AI-Powered Specialty Coffee Recommender System. This proposed system will simplify the coffee selection process for customers by leveraging AI and machine reasoning techniques to provide personalized, user-friendly recommendations.

The Community Coffee, in its pursuit to elevate coffee's status beyond a commoditized product, recognizes the challenge customers face when trying to navigate the diverse array of specialty coffee options. The complexity of coffee flavor profiles often leads to decision paralysis among customers and reduces operational efficiency. This challenge presents an opportunity for The Community Coffee to differentiate itself and enhance the customer experience through a personalized recommendation system.

The AI-Powered Specialty Coffee Recommender System is designed to streamline the coffee selection process by offering personalized coffee suggestions to customers. The system will use advanced machine reasoning techniques to analyze individual tastes, contexts, and preferences, converting the vast array of complex coffee flavor profiles into straightforward, easy-to-understand recommendations.

This system will aid our endeavor to enhance and refine these areas below:

- Enhanced Customer Experience: The user-friendly recommendation system will streamline the decision-making process for customers, thereby improving their overall shopping experience.
- Increased Efficiency: By simplifying the coffee selection process, we can increase operational efficiency and save time for baristas, leading to a more optimized service.
- Sales Growth: By delivering personalized recommendations, we can stimulate purchase decisions and upselling opportunities, contributing to revenue growth.
- Improved Customer Retention: A more customized and enjoyable shopping experience can boost customer loyalty and reduce churn rate.
- Market Differentiation: The AI-enabled recommender system provides a distinctive value proposition in a competitive market, fortifying our brand image.

Figure 3. The Community Coffee Shop

2.1 Problem Statement

In the burgeoning specialty coffee market, customers are increasingly overwhelmed by the complexity and diversity of coffee flavor profiles and options. This situation leads to decision fatigue, thereby limiting their ability to make informed and satisfactory purchasing decisions. Concurrently, baristas are spending significant time assisting customers in navigating the wide array of options, reducing their productivity and the overall efficiency of service.

Therefore, the challenge lies in effectively simplifying the coffee selection process by transforming the intricate flavor profiles into a user-friendly and intuitive system. There is a need to balance the provision of diverse coffee options with a simplified, personalized shopping experience that aligns with each customer's unique taste preferences and situations.

The current lack of a personalized recommendation system negatively impacts customer satisfaction, operational efficiency, and ultimately, sales potential. This presents an opportunity to develop a solution that leverages advanced AI technology to streamline the selection process, save time for baristas, and enhance the overall customer experience.

2.2 Market Research

2.2.1 Industry Overview

The global specialty coffee market has been on a consistent growth trajectory over the past few years, driven by factors like rising urbanization, increased disposable incomes, and changing lifestyle trends. There's a notable shift in consumer preferences towards high-quality, ethically sourced, and diverse coffee products.

2.2.2 Consumer Trends

Consumers increasingly value personalized experiences when making purchase decisions. A recent study shows that over 80% of customers are more likely to purchase from brands that offer personalized experiences. Further, with the rise of e-commerce and digital technologies, consumers are growing accustomed to recommendation systems used by brands like Amazon and Netflix.

2.2.3 Competitive Landscape

Several players in the coffee industry have begun to leverage digital tools to enhance customer experiences. However, the integration of advanced AI-based recommendation systems specifically for specialty coffee selection is still a largely untapped niche. This offers The Community Coffee a distinctive competitive advantage if implemented effectively.

2.2.4 AI-Based Recommendation Systems in Retail

AI-based recommendation systems have been successfully used in several industries, notably retail and entertainment. They have proven effective in boosting customer engagement, improving sales, and enhancing customer retention. Starbucks, a player in the coffee industry, has achieved success with its AI-based recommendation system, leading to increased customer spend and engagement.

2.2.5 Consumer Acceptance of AI-Based Systems

As AI becomes more commonplace in everyday applications, consumer acceptance of AI-based recommendation systems has improved. A study by PwC showed that 72% of business decision-makers believe that AI will be the business advantage of the future. Consumers are increasingly comfortable with AI recommendations and understand the benefits they offer in terms of personalization.

2.2.6 Conclusion

The market research indicates a promising environment for the introduction of an AI-based specialty coffee recommender system. There's a clear market trend towards personalized experiences, and AI-based recommendation systems have been successful in similar applications. As such, the proposed system is well-aligned with both industry trends and consumer preferences.

The key to successful implementation will be ensuring that the system truly understands and responds to the complexity of individual coffee preferences, ensuring data privacy, and effectively communicating the system's benefits to customers. With this in place, The Community Coffee can differentiate itself in a competitive market and capitalize on the opportunities for increased customer satisfaction and sales growth.

2.3 Proposed Solution

The proposed solution entails the development and implementation of an Advanced AI-Powered Specialty Coffee Recommender System, designed specifically to fit Community Coffee's product range and customer demographics. This system would apply machine reasoning techniques, including decision automation, knowledge discovery, data mining, and cognitive science, to offer personalized coffee recommendations to customers.

Features and Functionality:

- Moment-Based Recommendations: The system will propose coffee selections suitable for specific moments of the day - morning, afternoon, or anytime - factoring in the typical caffeine content suitable for these periods.
- Flavor Profiling: Users will be guided through a tiered flavor selection process. They will first choose from main flavors (nutty, chocolaty, spicy, sweet, floral, tea-like, herb-like, cereal, alcohol, and fruity), and then refine their selection based on sub-flavors. For example, if a user selects 'fruity', they will be asked to specify the type of fruit flavor they prefer.
- Coffee Type Selection: The system will offer customers the option to select their preferred coffee preparation method, like filter coffee or espresso
- Detailed Coffee Information: Once the system generates the top 3 recommendations, it will provide a detailed summary of each coffee, including the origin of the coffee bean, recommended consumption time and milk preference, and the detailed cupping notes.

The advanced AI-Powered Specialty Coffee Recommender System aims to simplify the coffee selection process for customers, promote efficiency, and enhance customer satisfaction, ultimately aligning with contemporary consumer trends towards personalized experiences. It will provide The Community Coffee with a competitive advantage by harnessing AI technology to drive business growth and improve customer service.

3. Knowledge Representation

3.1 Coffee Flavor Categorization Rule

The coffee flavor wheel is a valuable tool used by coffee professionals and enthusiasts alike to describe and categorize the complex flavors found in coffee. Just like the flavor wheel used in the world of wine, the coffee flavor wheel serves as a visual representation of the diverse range of tastes and aromas that can be experienced in a cup of coffee. It provides a standardized language for discussing and understanding the nuanced flavor profiles present in different types of coffee beans. By utilizing the coffee flavor wheel, individuals can better communicate their preferences, explore new flavors, and develop a deeper appreciation for the intricacies of the coffee they enjoy. Whether you're a barista, a coffee aficionado, or

simply curious about the fascinating world of coffee flavors, the coffee flavor wheel is an essential tool that unlocks a whole new realm of sensory exploration [4].

The coffee flavor wheel is divided into three distinct sections: the inner wheel, middle wheel, and outer wheel [Fig 4]. Each section represents a different level of flavor detail, allowing coffee professionals and enthusiasts to delve deeper into the intricacies of taste and aroma.

The inner wheel provides a broader overview of the coffee flavor profiles. It includes general descriptors that encompass a wide range of flavors such as fruity, floral, nutty, and sweet. These descriptors give a basic understanding of the dominant taste characteristics found in coffee and serve as a starting point for flavor exploration.

Moving to the middle wheel, we encounter more specific flavor categories that further define the taste experience. This section expands on the descriptors found in the inner wheel and includes sub-categories like citrus fruit, berries, dried fruits, brown sugar and more. These distinctions help to pinpoint the particular fruit flavors that can be detected in different coffees, offering a more nuanced understanding of their profiles.

The outer wheel of the flavor wheel is the most detailed and specific section. It delves into the finest nuances of flavor, providing intricate descriptors that capture the subtlest taste notes. This section includes descriptors like jasmine, black tea, blueberry, lemon, honey and many others. These descriptors allow for a more precise and detailed description of the sensory experience that coffee can offer.

By utilizing the inner, middle, and outer sections of the flavor wheel, individuals can navigate the complexity of coffee flavors with greater precision and clarity. It enables them to communicate and appreciate the full spectrum of taste and aroma found in different coffee beans, fostering a deeper understanding and enjoyment of the diverse world of coffee flavors.

Figure 4. Coffee flavor wheel

3.2 Introduction to Coffee Cupping Notes

Coffee cupping notes serve as a language to describe and evaluate the sensory attributes of coffee. Just as a sommelier assesses wine, coffee cupping allows professionals and enthusiasts to objectively analyze and appreciate the distinct flavors, aromas, and characteristics found in various coffees. Cupping notes are a vital part of the coffee tasting process, providing a standardized framework for assessing and comparing different coffee samples. During a cupping session, trained individuals slurp small spoonful of brewed coffee, allowing the liquid to cover their entire palate. As they taste the coffee, they pay close attention to the nuances of flavor, acidity, body, sweetness, and aftertaste. These sensory experiences are then meticulously recorded as cupping notes, which consist of descriptive language that captures the unique qualities of each coffee.

Cupping notes go beyond simple flavor descriptions by offering specific and detailed observations. They can include text descriptors such as fruity, floral, chocolatey, nutty, herbal, spicy, or even more distinctive notes like blueberry, jasmine, caramel, or cinnamon. Cupping notes also consider factors like the coffee's acidity, mouthfeel, and overall balance.

These carefully crafted cupping notes serve multiple purposes. They act as a guide for coffee professionals in selecting and sourcing beans, assisting in the creation of blends, and identifying exceptional coffees. Additionally, cupping notes help consumers make informed choices by providing a deeper understanding of the flavor profiles of different coffees. Whether you are a coffee industry professional or an enthusiastic coffee lover, cupping notes are an essential tool for unlocking the sensory journey and appreciating the remarkable diversity and complexity of the world's beloved beverage.

4. Data collection and preprocessing

Our team member, Gengwei, collected the raw coffee bean data by reaching out to the baristas at The Community Coffee shop. The data consists of the coffee bean name, varietal, producer, origin, processing method (including washing, fermentation, and drying), cupping notes (such as floral, brown sugar, milk chocolate…), package weight, price, and type of coffee (either filter or espresso). Additionally, Gengwei utilized her domain knowledge and considered the coffee bean processing method to label each coffee bean with its suitable consumption time and whether it is more suitable for filter or espresso preparation. This information helps in guiding users to select coffee beans that align with their preferred brewing methods and desired flavor profiles.

The most challenging aspect of data analysis was working with cupping notes, which contained complex and abstract flavor descriptions that corresponded to the outermost regions of the coffee flavor wheel. To address this challenge, we created a flavor dictionary that categorized the outermost flavors into a smaller number of main categories, making it easier to design the recommendation system. The flavor dictionary contains 10 main flavors, including nutty, chocolaty, spicy, sweet, floral, tea-like, herb-like, cereal, alcohol, and fruity. In the fruity category, flavors are further classified as citrus fruit, berry, stone fruit, tropical fruit, pome fruit, and dried fruit. For example: under the main category of “tea-like”, it contains a list of sub-flavors including: “darjeeling, ceylon, oolong, tencha, hibiscus, tisane, teh-o siew dai, matcha, roselle, earl grey”. We used a Python script to map the cupping notes

with the flavor dictionary and create a new column that transformed the cupping notes into main flavors. For example, the script can transform the cupping notes "Green plums, baked granny apples, light honey, red tea, marzipan" into the main flavor categories of "sweet, floral, tea-like, and citrus fruit". By doing this, we generated flavor features for each coffee bean, which not only facilitated the data analysis process but also converted abstract flavor descriptions into more comprehensible and intuitive flavor categories.

5. Recommender System Design

5.1 System Architecture Overview

The system architecture consists of two primary components: the Frontend and the Backend [Fig 5]. This division was chosen to enable parallel development of both components. The Frontend was developed alongside a simulated Backend, allowing for the creation of a user flow prototype.The Backend component of the system is responsible for hosting the computational logic of the recommendation system. It processes the user's quiz answers and generates the appropriate recommendations, which are then passed back to the Frontend for display.To facilitate communication between the Frontend and Backend, a JSON message containing the user's quiz answers is sent from the Frontend to the Backend. The Frontend then awaits a JSON message from the Backend containing the top three Coffee Recommendations. The architecture overview of the smart coffee recommendation system is depicted in Fig 6.

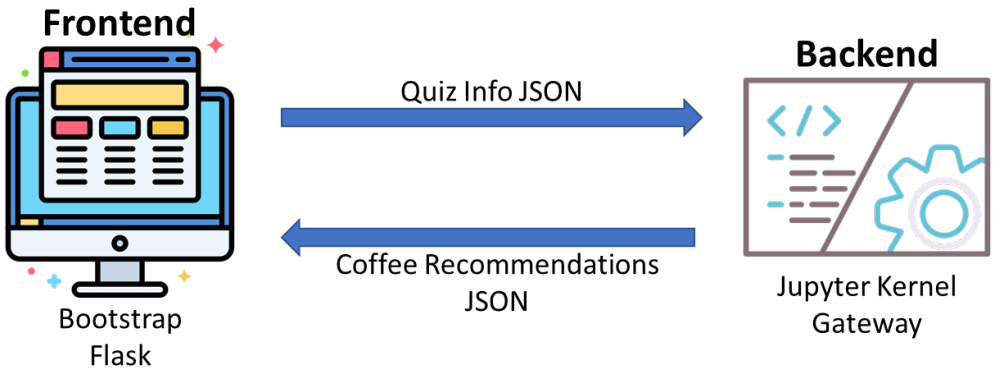

Figure 5. Schematic illustration of frontend and backend interaction

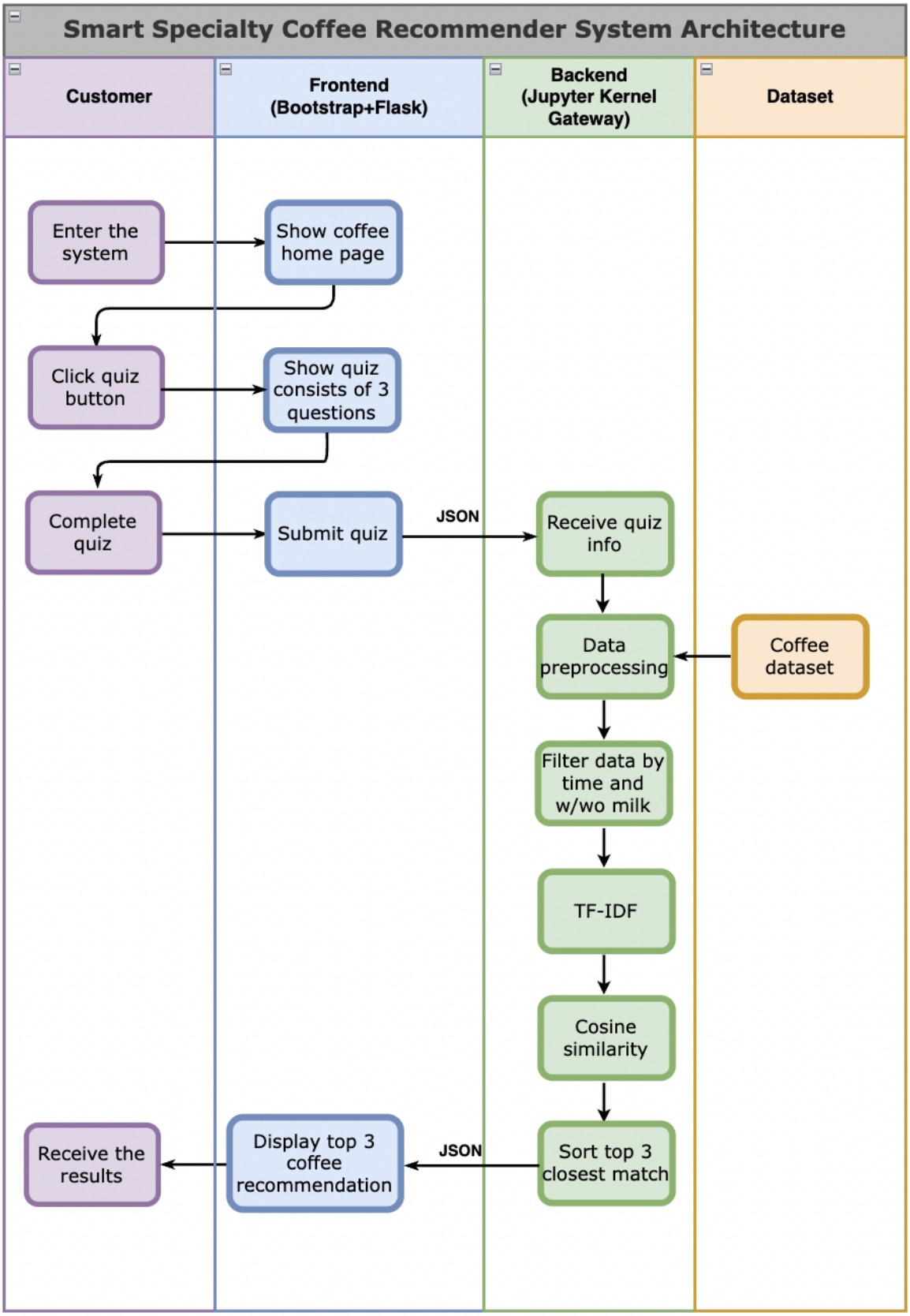

Figure 6. Overview of smart specialty coffee recommendation system architecture

5.2 Frontend Design

Bootstrap is a popular frontend development framework that helps web developers quickly and easily create responsive websites. It is a free and open-source toolkit that includes HTML, CSS, and JavaScript components. One of the main advantages of using Bootstrap is its responsiveness. Bootstrap's built-in grid system and responsive classes ensure that websites created with it are optimized for a wide range of screen sizes, from desktop computers to mobile devices. Additionally, Bootstrap provides numerous customizable components, including navigation menus, forms, buttons, and typography, which can be readily tailored to suit the specific requirements of a website. Our frontend web page has been developed using Bootstrap due to its inherent advantages, and its pre-designed templates and customizable components have significantly reduced the time needed to create high-quality websites.

The webpages are rendered and processed using Flask, a lightweight Python micro-framework. Flask allows us to keep the Frontend small and efficient. Additionally, the Flask framework facilitates communication with the Backend by enabling REST API-based interactions.

Our website consists of three pages, with the first serving as the home page that introduces our specialty coffee recommender [Fig 7-9]. By clicking on the "know your coffee today" button or clicking “quiz” button in the navigator bar, customers are directed to the second page, which features our quiz. The quiz consists of three questions that help us determine the customer's coffee preferences:

- Question 1: What is your favorite time to drink coffee? Is it morning, afternoon, or anytime?
- Question 2: What type of coffee do you prefer, filter (without milk) or espresso (with milk)?
- Question 3: Choose at least 3 flavors that appeal to you? There are 10 main flavors including nutty, chocolaty, spicy, sweet, floral, tea-like, herb-like, cereal, alcohol, and fruity. In the fruity category, flavors are further categorized as citrus fruit, berry, stone fruit, tropical fruit, pome fruit, and dried fruit.

Clicking on the "what's my coffee match" button directs customers to the product page, where they receive their top three recommended coffee beans along with detailed descriptions that include the coffee bean name, origin, recommended consumption time and milk preference, and the detailed cupping notes. If customers want to learn more about our store, they can click on the "store" page, which provides a detailed introduction to the different types of workshops that we offer to help customers gain a better understanding of specialty coffee.

Figure 7. Smart coffee recommendation website Home page and Store page.

Figure 8. Smart coffee recommendation quiz questions

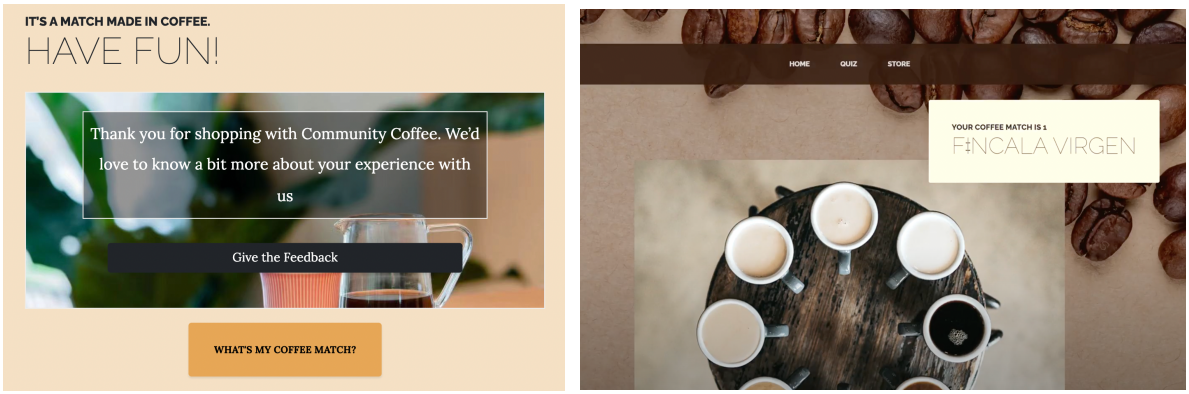

Figure 9. Clicking "What's my coffee match" button will display top 3 recommended coffee. An example of top recommended coffee

5.3 Backend Design

The Backend is a REST API built off a Jupyter Notebook, using the Jupyter Kernel Gateway Python service. This design choice offers flexibility for both the Frontend and Backend components, enabling seamless integration into various projects. By utilizing the Jupyter Kernel Gateway, the development process is accelerated, as it facilitates the conversion of data processing functionality from the initial Jupyter Notebook exploration to a fully functional REST API. This approach saves time that would have been spent on manually converting the algorithm into a different programming language to run on the server.

5.4 System Algorithm

The utilization of text similarity to build a recommendation system, which involves comparing customer input with a database, is a widely adopted approach for delivering personalized recommendations to users. One of the main advantages of utilizing text similarity recommendation systems is their ability to provide highly relevant recommendations based on a user's preferences and tastes. By analyzing the similarities between various coffee aromas and categorizing flavors based on the coffee wheel, the system can suggest coffee beans that are likely to be appealing to individual users. Another benefit is the relative ease of implementation and maintenance associated with similarity-based recommendation systems. The system can be designed to automatically analyze user choices and calculate similarity scores, which can be periodically updated with new data. For instance, when a new coffee bean is released, simply updating the database would suffice.

Our system algorithm utilizes the TfidfVectorizer to convert the cupping notes and customer's input into numerical vector representations, where each vector represents the presence or absence of specific words or terms. The cosine similarity is then applied to measure the similarity between these vector representations. By employing text similarity mining techniques, the code helps to recommend coffee options that closely match the customer's preferences, taking into account the time preference, milk preference, and the similarity of cupping notes to the provided keywords.

Cosine similarity is chosen as a metric for recommending coffee to customers because:

- It is a commonly used measure for comparing the similarity between two vectors in a high-dimensional space.
- Cosine similarity is particularly effective when comparing vector representations, as it measures the cosine of the angle between the vectors, providing a measure of similarity that is independent of the vector magnitude.
- It is unaffected by the magnitude of the vectors being compared, focusing on the relative orientation of the vectors rather than their lengths.
- Cosine similarity captures the contextual relevance between the customer's preferences (keywords) and the cupping notes, allowing for recommendations that closely align with the customer's preferences.
- Cosine similarity scores can be easily interpreted as a measure of similarity between 0 and 1, enabling the sorting and ranking of cupping notes based on their similarity to the customer's preferences.

The detialed steps of our system algorithm comprise the following:

1. Read the original coffee dataset from an Excel file using pandas.

2. Preprocess the dataset, if required, to ensure the data is in a suitable format for analysis.

3. Categorize the cupping notes in the dataset into main flavors. This can be achieved by mapping the cupping notes to flavors using a predefined mapping, as shown in the provided code.

4. Add a new column to the dataset called "Flavor" and populate it with the categorized flavors.

5. Save the updated dataset back to the Excel file.

6. Define the customer's input, which may include preferences such as time of consumption, milk preference, and keywords representing desired flavors.

7. Filter the coffee dataset based on the customer's preferences, such as time and milk preference, using pandas operations.

8. Create a TfidfVectorizer object to transform the cupping notes into vectorized representations.

9. Transform the customer's keyword string using the same vectorizer.

10. Compute the cosine similarity between the keyword vector and the vectorized cupping notes using the cosine_similarity function from sklearn.metrics.pairwise.

11. Add a new column to the filtered coffee dataset called "Cosine similarity" and populate it with the computed cosine similarity scores.

12. Sort the filtered dataset based on the cosine similarity scores in descending order.

13. Retrieve the top matching cupping notes and their corresponding information from the dataset, such as coffee name, recommend for, cupping notes, and origin.

14. Display the top 3 recommended coffee options to the user.

By following these steps, we can implement the cosine similarity method in a coffee recommendation system to provide personalized coffee recommendations based on user preferences and the similarity of cupping notes.

In our coffee recommendation system, we faced a significant challenge. Initially, we relied on cosine similarity to compare the flavor input provided by customers with the coffee flavor database. While this approach yielded high cosine similarity scores, it failed to consider important factors such as the recommended coffee drinking time and whether the coffee is

suitable for consumption with or without milk [Fig 10]. Figure 10 shows the initial attempt to capture only the customer's preferred flavor without considering their preferred time and milk preferences. For instance, when the customer inputs "tea, tropical fruit, sweet," the system provides the top 3 recommended coffees with a mix of suitable consumption times throughout the day. These considerations are vital as they differentiate distinguish between espresso-based drinks (served with milk) and filtered-based drinks (served without milk), which is a crucial aspect of customer preference. Additionally, the system's ability to provide moment-based recommendations is another key feature that enhances the overall user experience.

```
customer_input: tea, tropica fruit, sweet
Name: AQUIARES ESTATE
Recommend for: Espresso & Filter, morning
Cupping note: Green plums, baked granny apples, light honey, red tea, marzipan
Origin: Turrialba, Costa Rica
Varietal: Esperanza
Processing: Honey process
Cosine similarity score: 0.6567180107391848

customer_input: tea, tropica fruit, sweet
Name: SITIO VARGEM GRANDE #1
Recommend for: Espresso & Filter,morning, afternoon
Cupping note: caramel apples, whisked matcha, nougat, dried canberries, light custard
Origin: Minais gerais, brazil
Varietal: Yello catuai
Processing: 72 hours anaerobic fermentation, dried on patios & drying box
Cosine similarity score: 0.6539934297657487

customer_input: tea, tropica fruit, sweet
Name: KEBEDE MARO
Recommend for: Espresso & Filter, morning
Cupping note: Passionfruit, Cordial, Mango puree, Rooibos tea, golden kiwi
Origin: Konga, Yigracheffe
Varietal: Wolisho, Dega
Processing: Natural
Cosine similarity score: 0.6539934297657487
```

Figure 10. Illustration of the initial attempt to capture only the customer's preferred flavor without considering their preferred time and milk preferences.

To address this issue, we implemented an improvement by filtering the coffee dataset based on the user's preferences for time of consumption and whether they preferred their coffee with or without milk. This initial filtering step allowed us to narrow down the dataset to a more relevant subset. Next, we utilized the cosine similarity method on the filtered dataset to match the remaining cupping notes with the user's desired flavors. For instance, when the customer inputs "afternoon, espresso, tea, tropical fruit, sweet," the system provides the top 3 recommended coffees that are suitable for espresso-based coffee and suitable for afternoon consumption [Fig 11]. By incorporating the additional filtering step and then applying the

cosine similarity method, we have successfully incorporated features into our coffee recommendation system to capture the customer's preferred coffee drinking time, their preference for consumption with or without milk, and their desired coffee flavors. These additional parameters allow us to provide more accurate and tailored recommendations to each individual customer based on their specific preferences, which significantly enhances the overall user experience and satisfaction with our coffee recommendation system.

```
customer_input: afternoon, espresso, tea, tropica fruit, sweet
Name: SITIO VARGEM GRANDE #1
Recommend for: Espresso & Filter,morning, afternoon
Cupping note: caramel apples, whisked matcha, nougat, dried canberries, light custard
Origin: Minais gerais, brazil
Varietal: Yello catuai
Processing: 72 hours anaerobic fermentation, dried on patios & drying box
Cosine similarity score: 0.6539934297657487

customer_input: afternoon, espresso, tea, tropica fruit, sweet
Name: LOS IDOLOS
Recommend for: Espresso & Filter, morning,afternooon,anytime
Cupping note: Blood orange, Red apple, Toffee, darjeeling tea
Origin: Narino, colombia
Varietal: Caturra, castillo
Processing: Fully washed
Cosine similarity score: 0.6196178490982258

customer_input: afternoon, espresso, tea, tropica fruit, sweet
Name: FRUTA MADURA
Recommend for: Espresso & Filter, morning,afternooon,anytime
Cupping note: cooked apricots, toffee, apple cordial, ceylon tea, green grapes
Origin: Narino, colombia
Varietal: Castillo, colombia
Processing: Whole cherry carbonic maceration, pulped & dried in parabolic dryers
Cosine similarity score: 0.5881224174998241
```

Figure 11. Illustration of the improved algorithm that captures the customer's preferred flavor while also considering their preferred time and milk preferences.

5.5 Frontend and Backend Integration

The Frontend will send a JSON message with the following field to the Backend:

- Time of Day: Morning, Afternoon, or Anytime
- Coffee Type: Filter or Espresso
- Flavors: nutty, chocolaty, spicy, sweet, floral, tea-like, herb-like, cereal, alcohol, and/or fruity
- Fruity Flavors: citrus fruit, berry, stone fruit, tropical fruit, pome fruit, and/or dried fruit

Upon receiving the JSON message containing the aforementioned information, the Backend will convert it into a processable list for the data algorithm. Once the data algorithm has completed its processing, the Backend will encapsulate the results in another JSON message, including the top 3 coffee recommendations, which will be sent back to the Frontend for display.

5.6 System Validation

The validation process for the AI-Powered Specialty Coffee Recommender System is being conducted in phases to ensure a thorough evaluation and efficient error identification. Thus far, the following steps have been taken:

Unit Testing:

The first phase of validation involved unit testing of individual system components. This process confirmed the independent functionality of each feature, such as moment-based recommendations, flavor profiling, and coffee type selection.

Integration Testing:

Following unit testing, integration testing was performed. This phase involved testing the combined function of different system modules as a group. The objective was to ensure seamless interaction between system components and to confirm the system's performance when all parts function together.

System Testing:

The next step was system testing, during which the entire system was evaluated under conditions simulating real-world application. This testing phase allowed us to verify the system's compliance with business requirements and assess its usability, performance, compatibility, and security.

User Acceptance Testing (UAT) - Ongoing:

User Acceptance Testing is currently underway. In this phase, a select group of end-users are employing the system under real-world conditions. Their valuable feedback is being used to

identify practical issues that may not have been apparent during the theoretical testing phases.

Future Steps:

Following the completion of UAT, the system will be prepared for full rollout. Subsequent stages of validation will include post-implementation review and continuous improvement based on user feedback and system performance metrics. This continued validation effort will ensure that the system maintains its effectiveness and continues to deliver value to both Community Coffee and its customers.

6. Coffee Recommender System Commercial Values

The smart coffee recommender system offers substantial commercial value for coffee shops and businesses. Firstly, it enhances the overall customer experience by providing personalized coffee recommendations based on individual preferences. This tailored approach creates a sense of value and leads to increased customer satisfaction and loyalty. Additionally, the system drives sales and revenue by promoting specific coffee products or combinations based on customer preferences. By suggesting complementary items and upselling premium options, businesses can encourage customers to explore new offerings and make additional purchases, thereby boosting financial performance.

Furthermore, the recommender system presents cross-selling and upselling opportunities by identifying patterns and preferences among customers. By offering related products or accessories alongside coffee recommendations, coffee shops can seize cross-selling opportunities and promote higher-priced or premium items, increasing the average transaction value. Personalized recommendations also foster customer retention and repeat business. When customers consistently receive accurate and satisfying coffee suggestions, they are more likely to return to the coffee shop for future purchases, ensuring long-term customer loyalty.

In addition, the system generates valuable insights about customer preferences, enabling businesses to engage in targeted marketing and promotions. By understanding customer

tastes and preferences, coffee shops can tailor advertising campaigns and promotions that resonate with their target audience, leading to improved marketing return on investment. Moreover, implementing a coffee recommender system can provide a competitive advantage by offering a unique and personalized customer experience. This differentiation can attract new customers who are seeking personalized recommendations and a more engaging coffee-buying experience.

Lastly, the system generates valuable data about customer preferences, purchasing behavior, and trends, enabling coffee shops to make data-driven decisions regarding inventory management, menu design, product development, and overall business strategies. By leveraging this data, coffee shops can optimize their operations, make informed decisions, and stay ahead in the competitive market.

6.1 Advantages of System

The smart coffee recommendation system offers 4 advantages, including easier flavor selection for customers, timesaving for baristas, the ability to order in advance, and an improved user experience. Firstly, it simplifies the flavor selection process for customers, making it easier for them to choose their preferred coffee flavors using more understandable terms. This is especially helpful for customers who may not have in-depth knowledge of cupping notes or professional descriptions of coffee flavor. Secondly, our recommender system saves baristas time by eliminating the need for them to explain the detailed flavor of each bean. This allows them to focus on making coffee, especially filter coffee that takes more time. Thirdly, customers can access the system to order their coffee anytime and anywhere, allowing them to order in advance and save time by avoiding queues in the store. Lastly, our recommender quiz contains only four simple questions that take less than a minute for customers to answer. Besides, transforming the tedious flavor selection process into a simple quiz, it greatly enhances the user experience and turns it into a fun user experience.

6.2 Limitations of System

The smart coffee recommender system faces several limitations. Firstly, the system's data collection relies on a limited dataset from The Community Coffee shop, which constrains the

filtering criteria and restricts the range of recommendations. With a larger dataset, multiple recommendation algorithms could be introduced, providing a broader array of recommendations. Secondly, the system currently relies on a questionnaire-based user interface, which may lead to somewhat limited recommendation results and a restricted range of user diversity. This could potentially hinder the exploration of new and diverse coffee options. To address this limitation, additional system features can be implemented in the future, such as promoting new flavors or providing quick picks from this month's special offerings. These enhancements would offer users more opportunities to discover and try different coffee options. Lastly, privacy and security concerns arise from the collection and analysis of users' personal data and preferences, necessitating appropriate measures to address these concerns.

6.3 Future Outlook

Looking ahead, our coffee recommendation system holds immense potential for future development and expansion. While our system is currently implemented for a single local coffee shop, we envision extending its reach to encompass multiple coffee shops in Singapore, creating a comprehensive and dynamic coffee recommendation system for coffee enthusiasts and visitors alike.

In the future, our primary goal is to establish our recommendation system as a central hub for coffee lovers in Singapore. We plan to update the system on a monthly basis, ensuring that it remains up-to-date with the latest coffee offerings, trends, and seasonal specialties. By staying connected to the evolving coffee landscape, we can provide users with accurate and enticing recommendations that reflect the dynamic nature of the coffee scene in Singapore.

Expanding beyond a single coffee shop, we aim to collaborate with various coffee establishments throughout Singapore. By partnering with a wide range of coffee shops, from small local gems to renowned specialty coffee houses, we can offer a diverse selection of coffee experiences to our users. This collaborative approach will contribute to a thriving coffee community and foster a sense of exploration and discovery for both locals and visitors.

In addition to featuring a wide array of coffee options, our vision for the future includes integrating food and dessert pairings into our recommendation system. We believe that a delightful culinary experience is enhanced when carefully curated drinks complement the flavors and nuances of food and desserts. By incorporating this element, we can provide comprehensive recommendations that elevate the overall dining experience, allowing users to discover harmonious combinations of flavors and textures.

To bring this vision to life, we plan to create a comprehensive recommendation system map that showcases the diverse coffee shops and culinary delights across Singapore. This interactive map will serve as a go-to resource for coffee enthusiasts and food lovers, guiding them to exceptional coffee experiences and delightful gastronomic journeys. Users will be able to explore various neighborhoods, uncover hidden coffee gems, and discover the perfect pairing of drinks with delectable food and desserts.

As we embark on this journey, our mission is to foster a vibrant and thriving coffee culture in Singapore. We aim to connect coffee shops, customers, and culinary delights through our recommendation system, creating a sense of community and exploration within the local coffee scene. By constantly refining and expanding our services, we strive to enhance the enjoyment and exploration of coffee and culinary experiences in Singapore, ultimately establishing ourselves as the go-to platform for coffee and gastronomy enthusiasts.

With our future plans to regularly update the system, expand to multiple coffee shops, incorporate food and dessert pairings, and create an interactive recommendation system map, we are excited to contribute to the flourishing coffee scene in Singapore. We believe that our innovative approach will not only enhance the coffee experience for individuals but also promote the growth and appreciation of local coffee businesses. Through our dedication and commitment, we aspire to become a trusted and indispensable resource for coffee lovers, forging lasting connections between coffee, culinary delights, and the people of Singapore.

7. Conclusion

In conclusion, the development of an Advanced AI-Powered Specialty Coffee Recommender System presents a solution to the challenges faced by customers and baristas in the specialty coffee market, particularly The Community Coffee shop. The complexity and diversity of coffee flavor profiles often lead to decision fatigue among customers, while baristas spend significant time assisting customers in navigating the options. The proposed system aims to simplify the coffee selection process, balance diverse options with a personalized shopping experience, and enhance customer satisfaction.

The system's features and functionality include moment-based recommendations, flavor profiling, coffee type selection, and detailed coffee information. By leveraging advanced AI technology and machine reasoning techniques, the system offers personalized coffee recommendations based on the customer's unique taste preferences and situations. This not only streamlines the selection process but also saves time for baristas and improves the overall customer experience.

The smart coffee recommendation system offers several advantages, including easier flavor selection for customers, timesaving for baristas, the ability to order in advance, and an improved user experience. However, the system also faces limitations such as reliance on a limited dataset, narrow recommendation results, and privacy concerns. These limitations provide opportunities for future development and expansion.

Looking ahead, the goal is to establish the recommendation system as a central hub for coffee lovers in Singapore, extending its reach to multiple coffee shops and collaborating with various establishments. The system will be regularly updated to reflect the latest coffee offerings and trends. Additionally, the integration of food and dessert pairings into the recommendation system is envisioned to enhance the overall culinary experience.

Through these efforts, the coffee recommendation system aims to create a thriving coffee community, foster exploration and discovery, and provide users with comprehensive recommendations that elevate their coffee and dining experiences. With its potential for

growth and expansion, the system can play a significant role in shaping the coffee landscape in Singapore and beyond.

8. References

[1] "Colombian Milds-Other Milds differential tightens, I-CIP averages 170.03 US cents/lb in March 2023.," 2023. [Online]. Available: https://www.icocoffee.org/documents/cy2022-23/cmr-0323-e.pdf.

[2] "International Coffee Organization," 2021. [Online]. Available: http://www.ico.org/prices/new-consumption-table.pdf.

[3] "International Coffee Organization," 2021. [Online]. Available: http://www.ico.org/prices/po-production.pdf.

[4] M. Spencer, "Using Single Free Sorting and MultivariateExploratory Methods to Design a New CoffeeTaster's Flavor Wheel," *Journal of Food Science,* vol. 81, 2016.

9. Appendix

9.1 Appendix A: Project proposal

GRADUATE CERTIFICATE: Intelligent Reasoning Systems (IRS)
PRACTICE MODULE: Project Proposal

Date of proposal: 04 March 2023
Project Title: ISS Project – Intelligent specialty coffee recommender dedicated for the community coffee
Sponsor/Client: *(Name, Address, Telephone No. and Contact Name)* Institute of Systems Science (ISS) at 25 Heng Mui Keng Terrace, Singapore NATIONAL UNIVERSITY OF SINGAPORE (NUS) Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg
Background/Aims/Objectives: The proposed intelligent specialty coffee recommender applies advanced machine reasoning techniques including decision automation, knowledge discovery, data mining, and cognitive techniques to assist coffee drinkers in specialty coffee selection based on their preferred flavors or flavors that suit their moods/situations. It turns the complicated and non-intuitive coffee flavors into a user-friendly and

straightforward recommendation system, to ease the tedious coffee selection process for customers as well as saving time for baristas.

Requirements Overview:

- Recommendation ability
- System integration ability
- Data processing and analysis

Resource Requirements (please list Hardware, Software and any other resources)

Hardware proposed for consideration:

- Individual team members' laptops

Software proposed for consideration:

- Machine learning use cases: Orange3, PCA, cosine similarity, TF-IDF
- Frontend: React/Vue/Bootstrap/Flask
- Backend: Jupyter Kernel Gateway
- Application container: Docker

Number of Learner Interns required: (Please specify their tasks if possible)

a team of four to six project members (or individual work upon lecturer approval)

Gengwei Li: data collection, research on specialty coffee background, UI design, data analysis

Huiwen Tang: front end design using Bootstrap , system validation and verification

Ji Lin: documentation, video demo, system validation and verification

Jingyi Zhang: research on specialty coffee background, front end design using Vue3/Bootstrap, data analysis

Kelvin Fong: backend developer, system validation and verification

Methods and Standards:

Procedures	Objective	Key Activities
Requirement Gathering and Analysis	The team should meet with ISS to scope the details of project and ensure the achievement of business objectives.	1. Gather & Analyze Requirements 2. Define internal and External Design 3. Prioritize & Consolidate Requirements 4. Establish Functional Baseline
Technical Construction	To develop the source code in accordance to the design. To perform unit testing to ensure the quality before the components are integrated as a whole project	1. Setup Development Environment 2. Understand the System Context, Design 3. Perform Coding 4. Conduct Unit Testing
Integration Testing and acceptance testing	To ensure interface compatibility and confirm that the integrated system hardware and system software meets requirements and is ready for acceptance testing.	1. Prepare System Test Specifications 2. Prepare for Test Execution 3. Conduct System Integration Testing 4. Evaluate Testing 5. Establish Product Baseline
Acceptance Testing	To obtain ISS user acceptance that the system meets the requirements.	1. Plan for Acceptance Testing 2. Conduct Training for Acceptance Testing 3. Prepare for Acceptance Test Execution 4. ISS Evaluate Testing 5. Obtain Customer Acceptance Sign-off
Delivery	To deploy the system into production (ISS standalone server) environment.	1. Software must be packed by following ISS's standard 2. Deployment guideline must be provided in ISS production (ISS standalone server) format

		3. Production (ISS standalone server) support and troubleshooting process must be defined.

Team Formation & Registration

Team Name: Coffee Insider
Project Title (repeated): ISS Project – Intelligent specialty coffee recommender dedicated for the community coffee
System Name (if decided): Know your coffee
Team Member 1 Name: Jingyi Zhang
Team Member 1 Matriculation Number: A0269366M
Team Member 1 Contact (Mobile/Email): 86505893/e1112237@u.nus.edu
Team Member 2 Name: Lin Ji
Team Member 2 Matriculation Number:

A0269369H
Team Member 2 Contact (Mobile/Email): 90270668/e1112240@u.nus.edu
Team Member 3 Name: HuiWen Tang
Team Member 3 Matriculation Number: A0269361X
Team Member 3 Contact (Mobile/Email): 96110671/e1112232@u.nus.edu
Team Member 4 Name: Fong Zhi En Kelvin
Team Member 4 Matriculation Number: A0269365N
Team Member 4 Contact (Mobile/Email): 91284272/e1112236@u.nus.edu
Team Member 5 Name: Li Gengwei
Team Member 5 Matriculation Number:

A0269362W
Team Member 5 Contact (Mobile/Email): 96142511/e1112233@u.nus.edu
Team Member 6 Name:
Team Member 6 Matriculation Number:
Team Member 6 Contact (Mobile/Email):

For ISS Use Only		
Programme Name:	**Project No:**	**Learner Batch:**
Accepted/Rejected/KIV:		
Learners Assigned:		
Advisor Assigned: Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg		

9.2 Appendix B: Individual report

9.2.1 Jingyi Zhang's Individual Report

Name: Jingyi Zhang	**Student ID: A0269366M**

Personal contribution:

1. Ideation of the project
2. Conducted extensive research on the coffee flavor wheel and selection of system algorithm
3. Implemented a preprocessing method for the coffee data, which involved creating a flavor mapping dictionary to categorize outermost flavors into a smaller set of main categories. This approach streamlined the design of the recommendation system by transforming complex and abstract cupping notes into more manageable flavor categories.
4. Acquired proficiency in Bootstrap within a week and utilized it to design and beautify the frontend websites.
5. Developed a system algorithm capable of filtering the coffee dataset based on user preferences such as the time of consumption and the preference for coffee with or without milk. Then employed the cosine similarity method on the filtered dataset to match the remaining cupping notes with the user's desired flavors.
6. As the group leader, I effectively managed the project by assigning tasks based on each team member's expertise, monitoring their progress, helping team members when they encountered difficulties, and fostering communication and collaboration. I also undertook the responsibility of proofreading, editing, and formatting the group report.

Learning journey and outcome:

1. Theoretical knowledge

Through completing this graduate certificate program, I have acquired an extensive understanding of artificial intelligence (AI) and its various components. Specifically, I have gained knowledge in a wide range of machine learning algorithms, including decision

trees, K-nearest neighbors (KNN), fuzzy logic, informed and uninformed search, graph knowledge, and cosine similarity…etc. This comprehensive understanding of different machine learning techniques has equipped me with the ability to apply them effectively in practical scenarios.

Furthermore, I have developed a deep understanding of recommendation systems and their underlying principles. I learned collaborative filtering-based approaches, content-based techniques, knowledge-based methods, and community-based algorithms. This knowledge has provided me with valuable insights into the inner workings of recommendation systems, enabling me to design robust and efficient systems. This knowledge empowers me to create advanced solutions that leverage the capabilities of AI and deliver optimal results.

2. Tools and Platforms

In addition to the theoretical knowledge, this certificate program has provided various opportunities for practical application through workshop exercises. These exercises have allowed me to apply the theoretical concepts learned in real-world scenarios. I have gained proficiency in various tools and platforms, including: Ubuntu virtual machine, Orange, SQL, Neo4j, and google dialogflow. Additionally, I have become proficient in various Python libraries such as pandas, NumPy, spaCy, scikit-learn, and HMM learn. These libraries have provided me with a wide range of tools and functions for data manipulation, analysis, natural language processing, and machine learning tasks. By actively engaging in hands-on exercises with these tools and libraries, I have developed the skills necessary to tackle real-world problems effectively.

Application of Knowledge:

1. In the group project:

In our group project, I employed text similarity mining techniques, specifically TF-IDF and cosine similarity, to enhance our system's recommendation capabilities. By considering factors such as time preference, milk preference, and the similarity of cupping notes to the provided keywords, our system can suggest coffee options that closely align with the customer's preferences.

2. In the workplace:

As a research engineer in the field of science and engineering, I encounter various types of data in my daily work. Equipped with multiple machine learning algorithms, I am able to analyze this data efficiently. This skill allows me to extract valuable insights and make informed decisions based on the patterns and trends discovered in the data.

9.2.2 Fong Zhi En Kelvin's Individual Report

Name: Fong Zhi En Kelvin	**Student ID: A0269365N**

Personal contribution:

1. Exploration of Data, trying out one hot encoding, and some other basic processing
2. Exploration of different Frontend frameworks and Backend frameworks to use
3. Converting chosen frontend UI/UX into web server for use
4. Creating a REST API based backend on Jupyter Kernel Server for use
5. Integrating Frontend, Backend, and Data processing together
6. Provided instructions on how to setup the tech stack

Learning journey and outcome:

- Learned about how preparing data can be quite hard
- Learned that there are many different way to process the data, so many different similarity based algorithm to choose from, from Cosine to Jacard
- Learned how to use Flask to setup a server for Frontend use
- Learned how to use Jupyter Kernel Server to setup a REST API Backend
- Eventual outcome was better understanding of the similarity algorithms, and how they can be used for recommendation systems.
- Also had better understanding of how they can be integrated into actual systems

Application of Knowledge:

With any dataset, the first thing to do is to analyze the data, and pick up the features that can be used after understanding the data. Upon understanding the data, studies also need to be done on what algorithm can be used on the data, be it to predict outcomes for future data or for recommendation.

The final data processing may be a hybrid system, especially for more complicated data.

I do not work directly with data at my work, but may have to prepare systems to process the data. This program has helped me better understand how data can be processed, and why they are processed a certain way. This can help with how I prepare the systems for the data processing, like suggesting how the algorithms can be used, and using certain libraries instead that can help better with the processing.

I come back with a renewed appreciation for data engineers and data analysis. Picking up patterns, and learning about available data algorithms may not be that straightforward. It has been a mind opening experience learning and picking up new ways to think to solve the problems presented to us.

9.2.3 HuiWen Tang's Individual Report

Name: HuiWen Tang	**Student ID: A0269361X**

Personal contribution:

Following are my contributions in Coffee Recommendation System:

1. Determine system goals and requirements: Define the goals of the coffee recommender system, such as providing personalized coffee recommendations, helping users discover their favourite flavour coffee in Singapore Community Coffee Shop. Clarify the functional requirements, user interface requirements and data sources of the system. Search similar coffee websites to determine system goal and requirement, to get inspiration and references about the functionality, design, and user experience of other sites.
2. Webpage design and development: Design the user interface, including system introduction, coffee quiz display, recommendation result display, etc. Choose the appropriate front-end framework - Bootstrap and template for page development and layout design. Make sure the user interface is intuitive, easy to use, and beautiful.
3. Interface debugging：Establish a data interface with the front end to ensure normal data transmission and interaction. Work closely with the back-end team member for communication and debugging.
4. Project Report Preparation: Wrote up the user guide and project report for the segment mentioned above.

Learning journey and outcome:

- Learned about how to collect data and preprocess data,e.g. data clean, one-hot encoding, standardization or normalization, etc.
- Learned how to calculate similarity using different methods, e.g. cosine similarity, euclidean distance and jaccard similarity, etc.
- Learned how to use Bootstrap to design our recommendation system
- Learned how to use recommendation algorithms to model coffee data. Choose an appropriate algorithm based on needs, such as a content-based recommendation system model. Perform feature engineering to extract meaningful features, such as coffee attributes, user preferences, etc.
- Learned how to work closely with the team member, communicate and debug.

Application of Knowledge:

Recommendation systems are used everywhere in our lives, such as music and movie recommendation websites, travel and destination recommendation websites, shopping websites, real estate recommendations, etc.

If I have the opportunity to go to a large commercial software company in the future, I hope I can help the company build a product recommendation system, provide users with personalized recommendations, meet users' needs and preferences, increase their satisfaction with the platform, and Improve user loyalty. Increase sales and conversion rates, increase product exposure and variety.

Through the recommendation system, users can find the content they are interested in more quickly, and my company can better understand user needs and achieve business growth.

9.2.4 Gengwei Li's Individual Report

Name: Gengwei Li	**Student ID: A0269362W**

Personal contribution:

1. Idea Generation:

As a coffee lover, understanding the potential benefits of such a system, I proposed its development as a solution to enhance the user experience and provide personalized coffee recommendations even when they do not really know about coffee.

2. Domain Knowledge apply:

I apply my knowledge of coffee and related factors to design the system's functions effectively. By understanding the preferences and characteristics that influence coffee choices, I ensured the system's recommendations aligned with user expectations.

3. Database Creation:

Recognizing the significance of a robust database in the system's functioning, I took charge of creating and organizing the coffee-related data. By collecting comprehensive information on various coffee types, origins, processing and cupping notes, I ensured the system had a rich and diverse dataset for accurate recommendations.

4. Collaboration on Cosine Similarity Algorithm:

Working closely with my colleague, Jingyi, we collaborated on implementing the cosine similarity algorithm. By jointly exploring and experimenting with different approaches, we successfully integrated this algorithm into the recommendation system, enabling precise comparisons between coffee profiles and user preferences.

5. Business Value Outlook:

By analyzing market trends and consumer behavior, I identified the potential for growth and expansion of the Coffee Recommendation System.

6. Reporting and Documentation:

After finishing the project, I worked on the part of the report to keep our ideas and working process.

Learning journey and outcome:

1. Solving the "No Coffee Database Problem":

One of the challenges encountered in the project was the lack of an existing coffee database. I took the initiative to address this issue by creating a flexible database structure that allowed for easy updates and modifications. By tackling this problem, I gained experience in database management and learned to adapt to evolving project requirements.

2. Understanding and Applying Similarity Algorithms:

Through my involvement in the Coffee Recommendation System, I delved deeper into similarity algorithms such as cosine similarity, K-nearest neighbors (KNN), and decision trees. Applying the cosine similarity algorithm in this project allowed me to gain hands-on experience in measuring similarities between coffee profiles and user preferences. Furthermore, I expanded my knowledge of these algorithms and their applications, which proved valuable in other areas of work, such as using decision trees for predictive tasks in my office work.

3. Introduction to HTML Design:

During the project, I had the opportunity to explore basic HTML design principles. I learned how to make small design changes, such as modifying font styles and inserting icons, and gained insights into linking web pages. Though my understanding was based on examples, this introductory experience in HTML design expanded my skill set and provided a foundation for potential future web-based projects.

Application of Knowledge:

I applied my domain knowledge in the coffee area to make informed decisions and contribute effectively. I used the skills learned in our workshops, specifically in data cleaning, categorization, and filtering, to ensure the data used in the recommendation

system was accurate and relevant(For example, the raw data does not contain the recommendation for the drinking method, I did the categorization based on the processing method, origin and cupping notes). Additionally, I collaborated with Jingyi to experiment with various approaches using cosine similarity, ensuring that the results generated meaningful and personalized recommendations.

In my workplace, as a data analyst, I utilized the knowledge gained from studying different algorithms to identify the most suitable approach for a specific task. In particular, I discovered that using decision trees to create prediction models can be effective. I applied this understanding to analyze historical data and generate equations that capture relationships among multiple parameters. This approach was well-received by higher management, validating the practicality and value of the techniques I employed.

9.2.5 Lin Ji's Individual Report

Name: Ji Lin	**Student ID: A0269369H**

Personal contribution:

Following are my contributions in Coffee Recommendation System:

1. Euclidean Distance in Data Matching:

   I utilized this mathematical principle to measure the 'distance' between different coffee profiles based on various factors such as flavor, origin, and preferred time of consumption. By doing so, I provided another choice to make accurate and precise coffee recommendations tailored to individual customer preferences.

2. Front End User Interface Design Improvement:

   I also helped improve the design of the user interface for the recommendation system. My focus was on creating a user-friendly, intuitive, and visually appealing design that would simplify the coffee selection process for users. By incorporating animations and feedback systems, I ensured a seamless user experience that aligns with Community Coffee's commitment to superior customer service.

3. Video Production:

   Recognizing the power of visual storytelling, I contributed to the creation of an informative and engaging promotional video for the Coffee Recommendation System. This video explains the system's background, functionality, demonstrates its use, and highlights its benefits, making it an effective tool for user education and marketing.

4. Documentation Writing:

   I also helped with drafting comprehensive documentation for the Coffee Recommendation System. This included technical documentation review, problem statement, market Research, proposed solution etc..

Learning journey and outcome:

1. Learning Journey:

- Data Analysis and Algorithm Implementation: A major learning point was understanding and implementing the Euclidean distance method in data matching. This required deepening my knowledge of data analysis techniques and mathematical principles and how to apply them in the context of recommendation algorithms.

- User Interface Design: Another significant learning aspect was designing the front-end user interface. This task expanded my understanding of user experience (UX) principles, the importance of intuitive design, and the role of aesthetics in enhancing user engagement.

- Video Production: The creation of a promotional video for the system provided a valuable opportunity to hone my video production skills. This included storyboard creation, filming, editing, and the use of visual effects to convey information effectively.

- Technical Writing: Writing documentation for the system was a valuable exercise in effective technical communication. It improved my abilities to present complex information clearly and concisely, ensuring that it is easily understood by various audiences.

2. Outcome:

- Improved Technical Skills: My proficiency in data analysis, UX design, video production, and technical writing significantly improved, equipping me with a versatile skill set that can be applied to future projects.

- Enhanced Problem-Solving Skills: The various challenges encountered during the project development honed my problem-solving and critical thinking skills, making me more adept at troubleshooting and finding efficient solutions.

- Greater Understanding of User-Centered Design: Working on a project with a strong focus on user experience gave me a deeper understanding of user-centered design principles and their significance in the development of effective digital solutions.

- Expanded Knowledge of AI Systems: The hands-on experience with building a recommendation system deepened my understanding of AI systems, from theoretical concepts to practical application. This knowledge will be invaluable in any future work in the field of AI.

Application of Knowledge:

Leveraging my expertise in the coffee sector, I was able to make insightful judgments and contribute significantly to the project. I implemented the knowledge and skills acquired from our workshops, particularly those related to data cleansing, categorizing, and filtering, to guarantee the precision and relevance of data fed into the recommendation system. For instance, the raw data lacked guidelines for the drinking method; hence, I performed categorization based on parameters such as the processing method, origin, and cupping notes. Moreover, in partnership with Jingyi, I ventured into multiple methodologies utilizing cosine similarity, which ensured the generation of relevant and individualized coffee recommendations.

The knowledge and skills acquired from the Coffee Recommendation System project have also been incredibly beneficial in other areas of my work. For instance, in my role related to semiconductor equipment systems, I have been able to apply these skills to enhance system functionality, optimize data usage, and improve user interaction. The experience of using algorithms such as Euclidean distance and cosine similarity for creating recommendations has provided me with a greater understanding of data manipulation and matching techniques, which is invaluable in semiconductor equipment system work. Additionally, my experience with this project has significantly aided in my work with the

Employee Resource Group (ERG) mentoring program. I have utilized the same principles of data cleaning, categorization, and tailored recommendation generation to create more effective mentor-mentee matches. This cross-functional application of knowledge underscores the value of skills learned and their broader implications across diverse fields.

9.3 Appendix C: System User Guide

[1] To run the Backend
Using any Ubuntu based environment, or Ubuntu on WSL on Windows
$ pip install jupyter_kernel_gateway

Navigate to "BackEnd" folder

$ jupyter kernelgateway --KernelGatewayApp.api='kernel_gateway.notebook_http' --KernelGatewayApp.seed_uri='Backend.ipynb' --KernelGatewayApp.port=8899

[2] To run the Frontend
pip install flask

Navigate to "\FrontEnd\KnowYourCoffee" Folder

$ python main.py

Go to URL using web browser http://0.0.0.0:5000 or http://127.0.0.1:5000