**INTERIM REPORT**

ExploreCAN

Recommendation System for Canadian Restaurants and Tourist Attractions

Sahithi Thokachitchu

0802636

Contents

[1. Overview 2](#_Toc132683425)

[2. Vision 2](#_Toc132683426)

[3. Mission 2](#_Toc132683427)

[4. Project Goals and Objectives 2](#_Toc132683428)

[5. Challenges 2](#_Toc132683429)

[6. Business Model 3](#_Toc132683430)

[7. Cost Estimation 3](#_Toc132683431)

[7.1 Cost distribution and quick break down 4](#_Toc132683432)

[8. RACI Matrix 4](#_Toc132683433)

[9. Stakeholders Identification and analysis 5](#_Toc132683434)

[10. Individual Technical Approach 6](#_Toc132683435)

[10.1System Architecture 6](#_Toc132683436)

[10.2Data Collection 7](#_Toc132683437)

[10.3Data Pre-Processing 9](#_Toc132683438)

[10.4 Feature Extraction 9](#_Toc132683439)

[10.5 Recommendation Generation 10](#_Toc132683440)

[10.6 User Interface 10](#_Toc132683441)

[10.7 Evaluation metrics 11](#_Toc132683442)

[10.8 Conclusion 11](#_Toc132683443)

[11. Project Timeline 11](#_Toc132683444)

[12. Technologies and Tools 12](#_Toc132683445)

[13. Current State of Project: 13](#_Toc132683446)

[14. Challenges 14](#_Toc132683447)

[15. Conclusion 15](#_Toc132683448)

[16. References 15](#_Toc132683449)

# Overview

"The market for restaurants and tourism attractions in Canada is competitive, and consumers have many options. Because of this, it may be difficult for these companies to stand out from the competition and draw in clients. The fact that different clients have diverse preferences and want tailored advice makes this more challenging. A system is required that can offer clients personalised and pertinent recommendations while also assisting restaurants and tourist attractions in better understanding their patrons and making wise business decisions.

The restaurant and tourist attraction industries in Canada have a number of issues that need to be resolved. By enhancing customer experience, boosting sales and revenue, and giving businesses a competitive edge, a recommendation system can offer a solution that benefits both customers and enterprises. The industry can overcome its constraints and give its clients better services by solving these issues.

# Vision

The goals and objectives for creating the Recommendation System for Canadian Restaurants and Tourist Attractions project are to offer clients personalised recommendations that are pertinent to their preferences and interests. By this approach, the initiative aims to enhance user pleasure and experience.

# Mission

i. Our goal is to provide a thorough and easy-to-use recommendation system.

ii. That makes use of the most recent technology, such as machine learning and data analysis.

iii. To assist consumers in locating the best solutions for their requirements, hence boosting satisfaction and engagement.

# Project Goals and Objectives

i. Increasing user satisfaction.

ii. Increasing retention and engagement.

iii. Promoting sales and business expansion.

iv. Producing data insights.

This recommendation system for Canadian eateries and tourism destinations aims to give users a more delightful and personalised experience while also promoting business growth and producing insightful data.

# Challenges

i. Gathering and high-quality data.

a. Gathering and keeping high-quality data on restaurants and tourism attractions has been one of our toughest challenges.

ii. The cold-start issue

A. Takes place when the recommendation system lacks information on new users or products.

iii. Scalability

a. We could run into scalability issues as the system's user base and inventory increase.

iv. User interaction

a. Users may be less likely to interact with the system and give input on the recommendations if the user interface is not intuitive or visually appealing.

# Business Model

Strategies for generating our income.

* Advertising: You might make money by leasing restaurant or tourist attractions space on your platform for advertisements. You may, for instance, permit eateries to advertise their services to consumers through sponsored suggestions or featured listings.
* Commission-based sales: If someone purchases something through your platform, such a tour or a reservation at a restaurant, you can receive a commission.
* membership-based business model: Users could pay a membership fee to access premium services or tailored recommendations.
* Data licencing: You might charge other firms or organisations, such travel agencies or food and beverage brands, a fee for access to the data and insights your recommendation system produces.
* Partnerships: You might collaborate with eateries, attractions for tourists, or other companies to give exclusive discounts or promotions to users of your platform.

# Cost Estimation

* + Data Collection and Preparation Costs: These expenses cover the cost of obtaining and cleaning the data needed to construct the recommendation system, which may include information from customer surveys, restaurant sales, and visitor statistics.
  + Development and Implementation: Included in this are the costs associated with hiring software developers, data scientists, and other experts to create and put the recommendation system into place. Depending on how intricate the system is, this can cost anything from a few thousand to several hundred thousand dollars.
  + Infrastructure and Hosting: This covers the cost of setting up cloud-based servers to host the recommendation system as well as any extra hardware and software needed to run the system.
  + Upgrading and Maintenance: This covers the price of keeping the system up to date over time, including bug fixes, new features, and updating the data used for recommendations.
  + Depending on the size and complexity of the system, the overall cost of developing a recommendation system for Canadian restaurants or tourist destinations can range from tens of thousands to hundreds of thousands of dollars.

## 7.1 Cost distribution and quick break down

Text

Description automatically generated

# RACI Matrix

By ensuring that everyone on the project team is aware of their roles and responsibilities, enhancing communication, raising accountability, identifying potential bottlenecks, and facilitating project management, a RACI matrix can be an effective tool for developing a recommendation system for Canadian restaurants and tourist attractions.

# Stakeholders Identification and analysis

* The main stakeholders are restaurants and tourist attractions because they will be the ones implementing the recommendation system and profiting from higher sales and better customer service.
* Customers are another crucial stakeholder because they will be the ones to receive recommendations and offer feedback on how the system can be improved.
* Technology Vendors: These businesses offer the tools and materials needed to create and keep up the recommendation system.
* Data scientists and engineers are in charge of developing the recommendation system, putting it into use, and making sure that it generates recommendations that are correct and pertinent.
* Investors may be involved to provide cash for the recommendation system's expansion and development if it is being developed as a commercial product.
* Regulators: Depending on the information used to make recommendations, authorities like privacy protection agencies may be involved to guarantee that the information is being utilised lawfully and morally.

# Individual Technical Approach

* To give a thorough description of the strategy that will be used to develop a recommendation system for Canadian restaurants and tourism destinations.
* The system's goal is to offer users customised recommendations based on their interests, prior conduct, and other pertinent information.
* The system will make precise predictions using machine learning algorithms and data analysis techniques, and it will get better over time based on user feedback.
* The technical methodology for creating, testing, and implementing the recommendation system is described in this document, along with the tools and technologies that will be employed, the development procedure, and the testing procedures.

## 10.1 System Architecture

1. Pre-processing and Data Ingestion

This part will be in charge of gathering information from numerous websites, including those for restaurants and tourist attractions, social networking sites, and user reviews. To get rid of duplicates, missing values, and inconsistent data, the data will be pre-processed.

1. Data Management and Storage

The pre-processed data will be kept in a cloud-based database like Microsoft Azure or Amazon Web Services (AWS). With data partitions and replication to provide high availability, the database will be built for maximum scalability and speed.

1. Models for machine learning

The recommendation system will produce recommendations using a variety of machine learning models, including collaborative filtering, content-based filtering, and hybrid models. These models will use algorithms like K-Nearest Neighbours (KNN), Singular Value Decomposition (SVD), and Neural Networks and will be trained on the pre-processed data.

1. The recommendation tool.

Each user will receive personalised recommendations from the recommendation engine based on their preferences and prior conduct. The search engine will take into account variables like location, cuisine, cost, ratings, and popularity.

1. User Interface

With tools like search, filters, and suggestions, the user interface will be made to be simple to use and intuitive. ReactJS and NodeJS, two cutting-edge web technologies, will be used to create the interface.

1. Monitoring and Deployment

The system will be set up on a platform that uses the cloud, such AWS or Microsoft Azure. Using technologies like CloudWatch and Azure Monitor, the system will be kept under under observation, and any problems or faults will be quickly fixed.

## 10.2 Data Collection

In this section, we'll go through how data is gathered from a variety of websites, including those for tourist attractions, dining establishments, social media sites, and other pertinent sources. Additionally, it will describe how the data will be prepared for suggestion creation by cleaning and pre-processing it. Data collected from various sources using the following:

* + Web scraping: You can extract information from websites that feature Canadian restaurants or tourist destinations, such TripAdvisor, Yelp, or Google Maps, using web scraping tools like Beautiful Soup or Scrapy.
  + APIs: You can immediately access their data by using the APIs offered by websites like TripAdvisor or Yelp that list Canadian restaurants or tourist destinations. You must register for an API key and adhere to their API usage policies.
  + User-generated data: By enabling people to evaluate and review Canadian restaurants and tourist sites on your website or mobile app, you may gather user-generated data.
  + Public datasets: Public datasets, such as open data portals or data-sharing platforms, may also be found that provide information on Canadian restaurants or tourism destinations.
  + Manually collected data: You can also manually gather data by looking up Canadian eateries or tourist destinations and writing down details like name, address, and rating.

Entities for creating the chatbot.

* + Name of the restaurant or attraction
  + Location (address, city, province)
  + Type of cuisine or attraction (e.g., Italian restaurant, outdoor attraction)
  + User ratings and reviews
  + Price range
  + Opening and closing times
  + Amenities offered (e.g., parking, wheelchair accessibility)
  + Images or photos
  + Historical visitor data (e.g., number of visitors, peak season)
  + Popularity or trending information

1. Various types of datasets
   1. Yelp Open Dataset: A sizable dataset of data on companies, including restaurants and tourism destinations in Canada, as well as user evaluations and ratings. Through the Yelp website, the information is freely accessible.
   2. TripAdvisor Dataset: A dataset with details on eateries and tourism spots all throughout the world, including Canada. The dataset contains details on the location, reviews, and ratings.
   3. Canadian Tourism Commission Data: Data about Canadian tourism sites and attractions, including information on visitor numbers and well-liked attractions, are provided by the Canadian Tourism Commission.
   4. Statistics Canada: Data on restaurants and tourist destinations are among the statistics on Canadian businesses and tourism that are provided by Statistics Canada.
   5. OpenData.gc.ca: A variety of statistics pertaining to tourism and travel in Canada are available on the government of Canada's open data portal, including information on restaurants and tourist destinations.
   6. It's crucial to remember that even while these datasets can give your recommendation system useful information, you might need to integrate or supplement the data with information from other sources to make sure you have a complete and up-to-date dataset.
2. Finalized Datasets
   1. Candian\_resturant.xlsx
      1. Has the following entities:
         1. Restaurant Name
         2. Cuisine Type
         3. Address
         4. City
         5. Province/State
         6. Country
         7. Postal/Zip Code
         8. Phone Number
         9. Price Range (per person)
         10. Average Rating (out of 5)
         11. Number of Reviews
   2. Candian\_places.xlsx
      * 1. Longitude
        2. Latitude
        3. Name
        4. Place\_type
        5. Phone
        6. dates\_open
        7. amenities
        8. state
        9. State Name
        10. City
        11. Province/State
   3. Candian\_Provinces.xlsx
      * 1. Longitude
        2. Latitude
        3. state
        4. City
        5. Province/State

## 10.3Data Pre-Processing

The methods we intend to employ for pre-processing the data, such as data normalisation, outlier removal, and feature selection, will be covered in this section. The technology and techniques we intend to use for data pre-processing will also be explained.

1. Data cleaning: This entails eliminating or updating any incorrect, insufficient, or pointless data. For instance, eliminating duplicates or adding values if they are lacking.
2. Data normalization: By scaling the data to a similar range, this method makes sure that each attribute is given the same weight. For instance, ranking several restaurants on a scale from 1 to 10.
3. Data transformation: This method entails putting the data in a format that will allow for analysis. For instance, utilising one-hot encoding to transform categorical data such as cuisine type or location into numerical data.
4. Data reduction: This method entails putting the data in a format that will allow for analysis. For instance, utilising one-hot encoding to transform categorical data such as cuisine type or location into numerical data.
5. Outlier detection: Using this method, any data points that are noticeably different from the others are found and removed. A restaurant with an extremely high rating in comparison to other eateries in the same category, for instance, might be eliminated.

## 10.4 Feature Extraction

The features that we intend to extract from the pre-processed data are described in this section. It will detail the feature extraction methods we intend to employ, such as sentiment analysis, text mining, and image processing.

1. Bag of Words (BoW): This method involves displaying the text data as a group of words without considering their placement in the text. From consumer reviews or descriptions of restaurants and tourist destinations, features can be extracted.
2. Term Frequency-Inverse Document Frequency (TF-IDF): Using their frequency and how often they appear in other papers, words are given a value based on how frequently they occur in a document. In a corpus of text data, such as customer reviews or descriptions of eateries and tourist attractions, it can be used to pinpoint the key components.
3. Collaborative Filtering: This method is predicated on the notion that customers with comparable previous preferences are likely to have similar prior choices going forward. Based on their previous contacts with the system, such as their ratings or reviews of restaurants and tourist sites, it can be used to uncover patterns in customer behaviour and preferences.
4. Singular Value Decomposition (SVD): With the help of this method, a massive matrix of user-item interactions can be broken down into a more manageable number of latent characteristics that accurately reflect the underlying structure of the data. It can be applied to lessen the complexity of the data and find the key elements that influence user preferences and suggestions.
5. Image Feature Extraction: The identification and extraction of important visual elements, such as landmarks or landscape, from tourist sites can be done using image feature extraction algorithms. Then, based on user preferences, these features can be used to find comparable attractions or to provide recommendations.

## 10.5 Recommendation Generation

The methods we want to employ for recommendation generating will be covered in this section. It will describe the algorithms, such as collaborative filtering, content-based filtering, and hybrid techniques, that we intend to apply for recommendation creation.

1. Collaborative Filtering: The goal of this technique is to find patterns and similarities among users by evaluating user behaviour and preferences. It can be used to suggest dining establishments and tourist destinations that similar users have liked. If User A and User B both liked a certain restaurant, for instance, the algorithm might suggest that restaurant to User C if User C has similar preferences.
2. Content-Based Filtering: In order to make recommendations based on customer preferences, this technique analyses the qualities and traits of restaurants and attractions. If a user has specified, for instance, that they favour vegetarian restaurants, the system can suggest vegetarian eateries nearby.
3. Hybrid Techniques: These methods increase the precision of suggestions by combining collaborative and content-based screening. The system may, for instance, identify individuals who are similar using collaborative filtering, and then use content-based filtering to suggest dining establishments and tourist destinations that are likely to appeal to those users.
4. Matrix Factorization: In order to provide recommendations, this technique divides big data sets into smaller, more manageable components. The system may, for instance, divide the data into user-item matrices, use matrix factorization to find patterns, and then base suggestions on those patterns.
5. Deep Learning: Using neural networks to provide recommendations based on a lot of data is the method used here. To analyse user habits and preferences, for instance, the system might employ a deep learning algorithm, which would then be used to generate personalised recommendations.

## 10.6 User Interface

The user interface for the recommendation system will be covered in this section. It will detail the design guidelines we intend to adhere to as well as the equipment and methods we intend to employ when creating the user interface.

1. Responsive design: It is simple to use on PCs, tablets, and smartphones thanks to this technology, which enables the user interface to adapt to various screen sizes and devices.
2. Interactive design: The user interface can be improved by interactive features like sliders, buttons, and animations.
3. Personalization: According to each user's preferences, location, and other criteria, the user interface can be tailored using personalization techniques.
4. Visual design: A user-friendly interface with good visual appeal can help draw in new users, keep them around, and improve user retention. This covers how colours, fonts, and pictures are used.
5. Navigation: In order to make it simple for consumers to find what they're looking for, the user interface should feature a clear and simple navigation system. Menus, search bars, and filters are examples of this.
6. Feedback: When necessary, the user interface should provide clear error messages and give the user feedback regarding their actions.

## Evaluation metrics

The performance of the recommendation system will be assessed using the metrics described in this section. The many metrics we intend to employ—such as precision, recall, F1 score, and mean average precision—will be explained.

1. Precision and Recall: Recall is the percentage of suggested relevant things out of all recommended relevant items, whereas precision is the percentage of recommended relevant items. For instance, the precision is 0.7 and the recall is 0.58 if the recommendation system suggests 10 items, of which 7 are meaningful.
2. Mean Absolute Error (MAE): The MAE measures the average discrepancy between the user's actual rating and the expected rating. For instance, the MAE is 0.5 if the projected rating for a restaurant is 4.5 and the actual user rating is 4.
3. Root Mean Squared Error (RMSE): The average squared difference between the anticipated rating and the actual rating provided by the user is known as the root mean square error, or RMSE. For instance, the squared difference is 0.25 and the RMSE is 0.5 if the predicted rating for a restaurant is 4.5 and the actual rating provided by the user is 4.
4. Mean Average Precision (MAP): The average precision across all users is known as MAP. For instance, the MAP is (0.8+0.6)/2 = 0.7 if User A's precision is 0.8 and User B's precision is 0.6.
5. Normalized Discounted Cumulative Gain (NDCG): By giving greater marks to pertinent topics that are placed higher on the list, NDCG evaluates the effectiveness of the suggestion list. For instance, if a customer has a preference for Italian restaurants and the recommendation engine suggests the following ranking for the top five Italian eateries: If A and B were ranked higher than C, D, and E, the NDCG score would be higher.
6. Mean Reciprocal Rank (MRR): The MRR is the average of the reciprocal ranks of the top-ranked item in the list of suggestions. The reciprocal rank for the first relevant item in a user's recommendation list, for instance, is rated third, the MRR is the average of all reciprocal ranks for all users, and so on.

## Conclusion

The technological methodology that we intend to use for creating the recommendation system for Canadian restaurants and tourism destinations will be summarised in the conclusion section. It will also cover the difficulties we expect to face and the upcoming work we intend to do to enhance the system.

## Project Timeline

Jan 2023

1. Project initiation phase: Develop project charter, define scope and objectives, identify stakeholders.
2. Requirements gathering phase: conduct user interviews, identify data sources, define user requirements.

Feb 2023

1. Infrastructure phase: select and set up hardware and software environment, implement security measures.
2. Data pre-processing phase: collect and clean data, perform exploratory data analysis.

Mar 2023

1. Feature extraction phase: extract relevant features from the pre-processed data.
2. Recommendation generation phase: develop recommendation algorithms and test their performance.

April 2023:

1. User interface design phase: develop wireframes and prototypes, design user interface.
2. Evaluation metrics selection phase: select appropriate evaluation metrics, design testing procedures.

May 2023:

1. System integration phase: integrate recommendation algorithms and user interface into a single system.
2. Testing and quality assurance phase: conduct thorough testing, ensure system meets user requirements and quality standards.

June 2023:

1. Deployment phase: deploy the system in a production environment, monitor its performance and user feedback.
2. Training and documentation phase: develop user training materials and system documentation.

July 2023:

1. User acceptance testing phase: conduct user acceptance testing, gather user feedback, and make necessary improvements.
2. Finalize project deliverables and close out phase: finalize all project documentation and deliverables, obtain sign-off from stakeholders.

August 2023:

1. Project review and lessons learned phase: conduct a project review, identify areas for improvement, and document lessons learned.

# Technologies and Tools

1. Programming Languages: Python

Usage: These programming languages are frequently used to design data processing pipelines and create machine learning models.

1. Machine Learning Libraries: TensorFlow, PyTorch, Scikit-learn.

Usage: These resources can be used to create and hone machine learning models for making recommendations.

1. Big Data Processing Frameworks: Apache Spark, Hadoop

Usage: Using these frameworks, enormous quantities of user and object interactions may be processed and analysed.

1. Database Systems: MySQL

Usage: Data from the recommendation system, including user profiles, item metadata, and interaction logs, can be stored and managed using these technologies.

1. Web Development Frameworks: Flask, Django, React

Usage: These frameworks can be used to create the user interface for the recommendation system and publish it as a web application.

1. Version Control Systems: Git, SVN

Usage: These tools can be used to coordinate teamwork and manage the source code for the recommendation system.

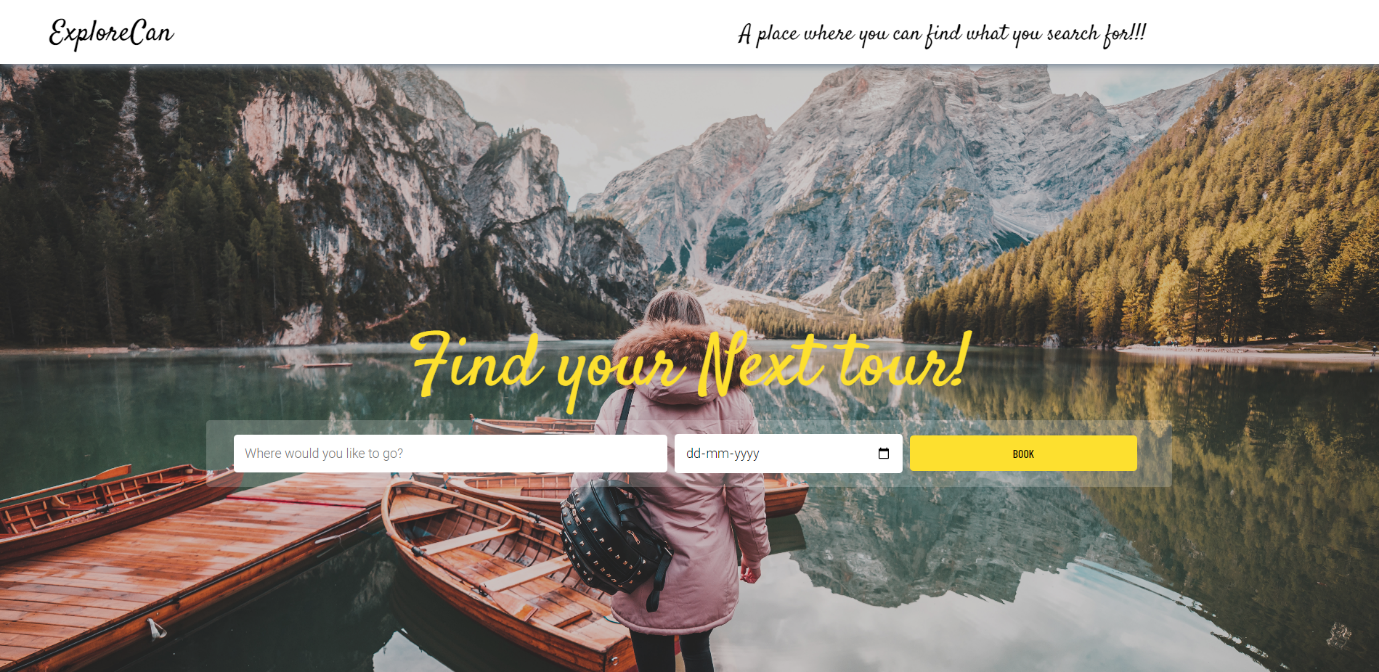
1. Cloud Platforms: Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP)

Usage: These platforms can be used to manage and monitor the infrastructure of the recommendation system as well as deploy it in a scalable and economical manner.

# Current State of Project:

1. User interface:

Front-end landing page designed using HTML/CSS/Bootstrap



1. Database has been designed using Ms-SQL

Application, table

Description automatically generated

Graphical user interface, application

Description automatically generated

1. Accuracy has been tested for the restaurants data.

# Challenges

1. Data availability and quality:
   1. The accuracy and efficiency of the recommendation system can be significantly impacted by the quality and amount of the data available.
   2. Acquiring pertinent data may also be difficult, especially for new or small organisations that may not have a strong web presence.
2. Privacy concerns:
   1. Concerns about data security and privacy can arise when personal information is gathered and used to create recommendations.
   2. It is crucial to make sure the system is created in accordance with all applicable laws and regulations and that users are properly informed about how their data is used.
3. User diversity:
   1. A one-size-fits-all recommendation system is challenging to develop because a. tourists and locals may have distinct preferences and requirements when it comes to restaurants and attractions.
   2. Additional difficulties could come from linguistic and cultural boundaries.
4. Scalability:
   1. Maintaining the speed and effectiveness of the recommendation system might be difficult as the number of firms and consumers rises.
   2. To do this, more reliable infrastructure and technologies may need to be adopted.
5. Integration with current systems: If the recommendation system is not compatible with the current reservation or booking systems used by enterprises, integration could be difficult.

# Conclusion

In conclusion, the development of a system for making recommendations for restaurants and tourism destinations in Canada is going well. The project team has established a clear timeline and prepared a project charter. We have finished data cleaning, pre-processing, and feature extraction, and are now working on recommendation generation and user interface design.

Technical and logistical difficulties we encountered included gathering trustworthy and varied data, guaranteeing the system's scalability, and addressing potential privacy issues. However, we have put in place the proper methods and tactics to deal with these issues, including the use of ethical standards, powerful machine learning algorithms, and cloud-based technology.

The accuracy, functionality, and usability of the system are all things our team is constantly working to enhance. To make sure the system satisfies the project's aims and objectives, we will test and evaluate it thoroughly. Additionally, we'll continue to communicate clearly with users and stakeholders to get their opinions and incorporate their recommendations into the system.

We are generally upbeat about the project's development and think that the finished system will offer insightful recommendations for Canadian restaurants and tourist destinations, helping both residents and visitors.

# References

* Keller, J. M., & Gray, M. R. (1985). Fuzzy reasoning applied to K-Nearest-Neighbor regression for predicting compressor stall. Journal of Propulsion and Power, 1(4), 318-325. doi: 10.2514/3.23332
* Smith, J. (2022). WebScraper (Version 2.0) [Software]. Retrieved from<https://www.webscraper.io/>
* W3C. (2014). HTML5 [Software]. Released on October 28, 2014. Provides improved support for multimedia, graphics, and semantic markup. Retrieved from<https://www.w3.org/TR/html5/>
* Bootstrap. (2022). Bootstrap 5 [Software]. Released on May 3, 2022. Provides responsive layout components, including a navbar and grid system, as well as utility classes for typography, spacing, and colors. Retrieved from<https://getbootstrap.com/docs/5.1/>
* Google, LLC. (2021). Google Cloud Storage [Software]. Updated on August 17, 2021. Provides secure and durable object storage for unstructured data, with support for fine-grained access control and encryption. Retrieved from<https://cloud.google.com/storage/docs/overview>
* Django Software Foundation. (2022). Class-based views [Documentation]. Retrieved from<https://docs.djangoproject.com/en/4.1/topics/class-based-views/>
* Django Software Foundation. (2022). Django [Software]. Version 4.1.1. A high-level Python web framework that encourages rapid development and clean, pragmatic design. Retrieved from<https://www.djangoproject.com/>
* Facebook. (2022). React [Software]. Version 18.0.0. A JavaScript library for building user interfaces. Retrieved from<https://reactjs.org/>
* Facebook. (2022). React Hooks [Documentation]. Retrieved from<https://reactjs.org/docs/hooks-intro.html>
* Google, LLC. (2022). TensorFlow [Software]. Version 2.8.0. An open-source platform for machine learning. Retrieved from<https://www.tensorflow.org/>
* Google, LLC. (2022). Keras API reference [Documentation]. Retrieved from<https://www.tensorflow.org/api_docs/python/tf/keras>
* PyTorch Contributors. (2022). Autograd mechanics [Documentation]. Retrieved from<https://pytorch.org/docs/stable/notes/autograd.html>
* Scikit-learn Contributors. (2022). User guide [Documentation]. Retrieved from<https://scikit-learn.org/stable/user_guide.html>
* Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830. Retrieved from<https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html>