INSY660 REPORT

CHATBOT

TEAM 4

HONGYI ZHAN YANHUAN HUANG GE GAO

CONTENT

ABOUT CONCORD

Concord, a gateway to seamless journeys and comprehensive solutions based in Montreal, offers inbound Canada and outbound Asia adventures, plus transportation and conference planning services.

Established in 1995, it emerged as a recognized and accredited inbound tour operator and earned Approved Destination Status (ADS) from the Tourism Industry Association of Canada.

CURRENT PROBLEM

Issue 1

While customers can make friends and build connections during travel after register for travel package, the agency's current system assigns travel companions on a **first-come, first-served** basis. With a chatbot implementation, potential travelers can independently **choose like-minded companions**, enhancing their journey with newfound friends.

Issue 2

Concord's present website spotlights travel packages, does **not fully address destination exploration**. By integrating a chatbot offering **tailored** travel tips and destination insights, users can engage more meaningfully. This upgrade is positioned to draw a larger user base and drive customer growth via an **enriched user experience.**

Issue 3

While the prices of travel packages remain constant, airfare and other expenses often vary. The website's current service **lacks** the capability to **offer related cost estimates**. Through the implementation of the new chatbot, customers can receive assistance in budget planning, enabling them to embark on their journeys at more advantageous prices.

GOAL

CUSTOMIZATION

DATASETS

Customer
Name
Gender
Age
Destination_Category
Temperature
Budget
Purpose
Num_of_People
Families
Duration (in days)
Travel Months
Transportation
Hobby
Language

Travel Cost
Purpose
Start_Date
End_Date
Destination
Airefare
Other_Transport
Lodging
Meals
Other_Expenses
Total
Additional_Comment

Airfare
Flight
Price_day1
Price_day2
...
Price_day30

Destination
Destination_Name
Temperature
Budget
Purpose
Language
Family_Friendly

Travel Tips
Destination
Introduction
Summary Tips
Transportation
Culture
Language
Weather
Photography Spots
Shopping
Tipping
Internet

DATASETS

- **Customer**
  - 10,000 customer records with randomly generated names, age between 18 and 65, and intended travel duration within 15 days
  - 10 destination preferences; 5 hobby categories; 5 transport preferences
- **Destination**
  - A compilation of 22 travel cities within Canada
  - categorized based on:temperature, consumption level (budget), prevalent travel purposes, predominant language, and family-friendliness.
- **Travel Cost**
  - a subset of the open.canada.ca public dataset, features records have only one destination city
- **Airfare**
  - simulates economy-class straight flight costs from the cities in *destination dataset* to Toronto
  - price range obtained from aircanada.ca with a 10-day timeframe, mimic airfare prices for a 30-day period
- **Travel Tips**
  - Python dictionary with travel trips and introductions generated by ChatGPT for the same 22 cities featured in the destination dataset

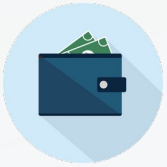
SCENARIOS

Our chatbot aims to support the customized travel service for Concord Travel Agency.
We think that the meaning of travel is **exploration and discovery**, **culture enrichment**, and **building relationships and connections**.
And so, we develop four scenarios to fulfill these travel purposes to improve customization, and these four scenarios are **all new scenario**.

Scenario 1: Recommendation on destination
In this scenario, we would like to compare the similarity between customer preference and city characteristics. We will recommend the destination city with the highest similarity to our customer. By doing so, we allow customers to search for any places that they may be interested in, and so to fulfill the purpose of exploration and discovery.

Scenario 2: Travel tips
In this scenario, we would like to provide some travel tips to the customer based on their selections. For now, we offer users to search for general tips, transportation, culture, language, weather, photography spots, shopping, tipping, and the internet. This allows our chatbot to develop following the purpose of culture enrichment.

Scenario 3: Find friends for group travel
This scenario would help the customer to connect with other travelers who have similar travel habits. By doing so, it allows the customer to form their own group travel and minimize the conflicts during the trip, and so fulfill the purpose of building relationships and connections.



Scenario 4: Estimated travel costs
After we implement the scenarios to fulfill the three main purposes of travel, we also want to give an estimated travel cost for our customers to help them plan their budgets. We apply ARIMA, a time series model, to find the future transportation cost, and add this cost with the average local expenses to compute the predicted budget.

SCENARIO 1

Personalized Destination Recommendation System

RATIONALE

In response to the evolving landscape of personalized travel experiences, our initiative seeks to revolutionize travel recommendations by leveraging the synergy between customer preferences and city characteristics. By quantifying the likeness between individual preferences and diverse urban traits, we aim to empower travelers to embark on journeys tailored precisely to their interests, while simultaneously nurturing a sense of adventure and discovery. By transcending conventional suggestions and introducing explorers to destinations that resonate on a personal level, our approach redefines travel, ensuring meaningful and fulfilling experiences for every individual seeking to uncover the world's treasures.

FUNCTIONALITIES

Utilizing the Destination dataset, which comprehensively encompasses destination attributes such as local temperature, expense levels, city categorization, predominant spoken language, and family suitability, our approach revolves around eliciting and processing customer information and preferences. By subsequently calculating the similarity between customer preferences and the aforementioned destination attributes, we employ the Euclidean distance metric. This choice is predicated on our discrete, small-valued feature set, rendering the application of cosine similarity unnecessary. The culminating step involves suggesting the destination city that exhibits the smallest L2-norm when contrasted with the customer vector. Through this methodology, we strive to provide tailored travel recommendations that harmonize with individual preferences and characteristics, enhancing the travel experience for our valued users.

SCENARIO 2

Local Travel Tips System

RATIONALE

In this context, our aim is to enrich users' cultural experiences through the provision of tailored travel tips aligned with their preferences. By offering a range of categories such as general advice, transportation, cultural insights, language assistance, weather guidance, photography hotspots, shopping recommendations, tipping etiquette, and internet accessibility, our chatbot is designed to empower travelers with comprehensive and personalized information. This approach facilitates not only smoother journeys but also deeper cultural immersion, as users are equipped with insights that transcend the typical tourist experience. By bridging the gap between practical advice and cultural enrichment, our chatbot serves as an invaluable companion for explorers seeking to engage authentically with the diverse world around them.

FUNCTIONALITIES

In this scenario, we have adopted a nested dictionaries structure, associating each destination with a dictionary encompassing nine distinct categories of travel tips. Upon user interaction, we solicit destination input along with the specific tip category of interest. Subsequently, we utilize this information to query the nested dictionary, extracting and displaying the relevant tips to the user's screen. This implementation facilitates seamless access to a tailored set of tips, enhancing user experience and aiding their travel preparations.

SCENARIO 3

TravelTribe Intelligent Grouping System

The TravelTribe Intelligent Grouping System is a forward-thinking solution designed to optimize group travel experiences by leveraging advanced technology and data analysis. It offers efficient itinerary planning, personalized experiences, and minimized coordination efforts, enhancing the overall satisfaction of travelers. Notably, the system facilitates connections among participants who share similar interests, fostering pre-trip engagement, tailored group dynamics, and meaningful post-trip connections. By seamlessly integrating technology, this system revolutionizes group travel, making it not only about the destination but also about forming lasting relationships with like-minded individuals.

RATIONALE

Personalized Group Formation: It utilizes advanced algorithms, specifically K-means clustering, to intelligently group participants based on their personal preferences across various aspects (weather, group size, age, travel tool, etc.) of travel. This enables the creation of well-matched travel companionships. The system also employs a range of diverse visualization plots to effectively present group information, facilitating collaborative planning and decision-making. Importantly, privacy is maintained through robust data protection measures, ensuring the confidentiality of personal information. This comprehensive solution not only enhances group travel coordination through personalized matching but also empowers participants with insightful visualizations while prioritizing their privacy and security.

FUNCTIONALITIES

SCENARIO 4

TravelPricisio Predictive Expense Navigator

RATIONALE

The TravelPricisio Predictive Expense Navigator is a system developed to empower travelers with accurate and predictive information about their travel expenses. It enables travelers to plan, budget, and make informed decisions by forecasting costs related to accommodations, transportation, activities, and more. This system aims to enhance travel experiences by promoting responsible spending, avoiding unexpected financial surprises, and streamlining the decision-making process. By offering personalized recommendations, transparent pricing insights, and data-driven predictions, the system contributes to more satisfying and well-managed travel journeys.

FUNCTIONALITIES

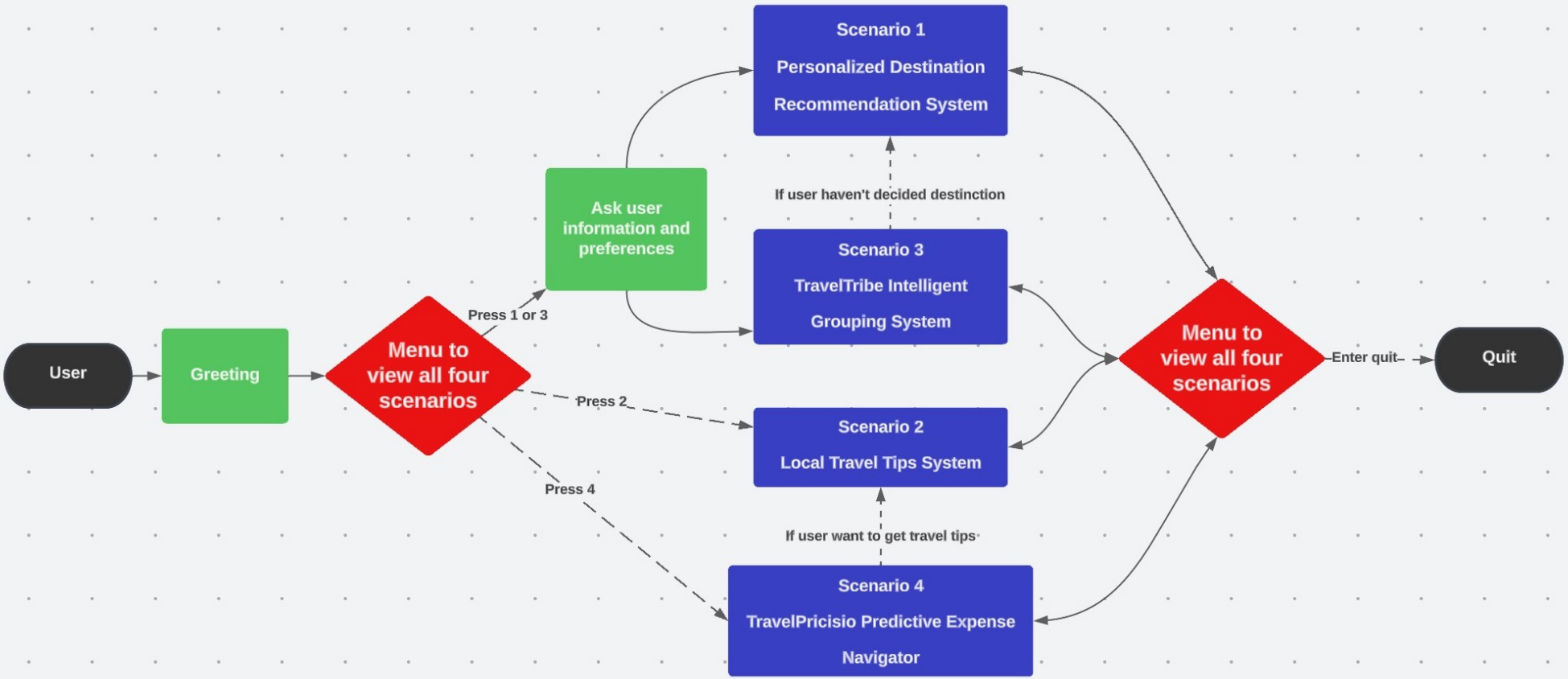
Expense Prediction: It can combines sophisticated time series analysis using ARIMA models for airfare prediction and the visualization power of pie charts to create a robust platform. Leveraging government data from Canada, the system provides travelers with accurate forecasts of airfare costs, considering historical trends and seasonality. The pie chart visualization enhances financial awareness, offering a clear and intuitive breakdown of budget allocation across various travel aspects such as accommodations, transportation, meals, and activities. By integrating these components, the system empowers travelers to make strategic booking decisions and effectively manage their expenses while providing a reliable and visually engaging tool for budget planning.

DESIGN STRUCTURE

In this project, we write a class Chatbot to perform all the scenarios. We would talk about the functions we use and how they cooperate together below.

- **__init__(self)**: Constructor that stores all the customer input information and preferences.
- **greeting(self)**: Greeting with users, allows user to input the number of option (scenario) they want the chatbot to display. We handle invalid input error by writing a while loop, and we write a while loop that allows user to continue requesting until they enter "quit". We also allow users to view the options menu if they forget the serial number of scenarios.
- **get_info(self)**: Ask for user's information and preferences. We handle invalid input error by writing while loops for each input information. This function will be called only once during the whole execution since scenario 1 and scenario 3 share the customer information, it is redundant to continuously ask for user's preference.
- **recommendation(self)**: Perform scenario 1. After performing the scenario, user will return to greeting page.
- **travel_tips(self):** Perform scenario 2. After performing the scenario, user will return to greeting page.
- **group(self)**: Perform scenario 3. After performing the scenario, user will return to greeting page.
- **mapcity(self)**: This is a helper function for scenario 4 that maps the airport with the city.
- **estimate(self)**: Perform scenario 4. After performing the scenario, user will return to greeting page.

DESIGN STRUCTURE



DATA ANALYTICS LIBRARY

Scikit-Learn

- Scikit-Learn (sklearn) is a popular machine learning library in Python that provides tools for various tasks, including clustering. K-means clustering is a commonly used algorithm to partition a dataset into a specified number of clusters.
- When planning group travel, it's important to consider the preferences and interests of all group members. Clustering helps identify destinations that are likely to be enjoyed by the entire group. This minimizes conflicts and disappointments arising from choosing a destination that only caters to a subset of the group.

Statsmodel

- The statsmodels library in Python is a powerful tool for statistical modeling and analysis, including time series analysis. One of its components is the ARIMA (AutoRegressive Integrated Moving Average) model, which is widely used for time series forecasting.
- In the context of transportation costs and travel budget forecasting, statsmodel can help provide insights into future spending patterns, allowing for better budget allocation, resource planning, and decision-making for travel-related activities. It enables data-driven predictions that are essential for optimizing travel experiences while staying within budget constraints.

DATA ANALYTICS LIBRARY

collections

- Calculating Frequency Distribution: The Counter is great for calculating the frequency distribution of elements in a dataset.
- Identifying Common Elements: most_common() helps identify the most frequent elements, which can be valuable in summarizing data.

IPython.display

- The IPython.display library is especially useful for creating rich and interactive content within Jupyter Notebooks or other IPython environments.
- Displaying External HTML Files: Display the content of an external HTML file such as picture of the landscape using the HTML class.

Matplotlib & Wordcloud

- Both WordCloud and Matplotlib are popular Python libraries used for data visualization. They can be used together to illustrate K-means clustering results for group travel.
- Using WordClouds, scatter plots, and other visualizations in combination with K-means clustering results can greatly enhance the understanding of travel destination clusters. Visualizations make it easier for travelers to make decisions aligned with their preferences and enable travel planners to effectively communicate recommendations.

matplotlib

ADDITIONAL LIBRARY

For future suggestion on the additional Python library, since our chatbot lacks attractive UI design and a more intelligent language model to response user's request like a human, we would like to suggest the following libraries for further improvements.

Tkinter

- The tkinter package enhances the user interface of applications, providing native-looking elements like buttons and menus for improved interactivity. The integration of the tkinter package into a travel chatbot vastly enhances the user experience by providing a user-friendly, consistent, and interactive interface. By leveraging tkinter's native design, interactive components, and layout management, the travel chatbot becomes a powerful tool for travelers to access tailored recommendations, plan itineraries, and engage in informative and engaging interactions.

TensorFlow / Pytorch / OPENAI API

- Integrating technologies like TensorFlow or PyTorch for natural language understanding and the OpenAI API for advanced language generation can greatly enhance a travel chatbot. These technologies enable the chatbot to comprehend user intent and context more accurately, generate more natural and engaging responses, and offer personalized travel recommendations and information. With these advancements, the chatbot becomes a powerful tool for users to plan trips, receive creative suggestions, and engage in informative conversations, ultimately delivering a more satisfying and user-centric travel planning experience.

SUGGESTIONS FOR CONCORD

- **Data Collection and Feedback Utilization**
  - Establish a database to store customer interactions and feedback garnered from the chatbot. This will not only inform campaign strategies for targeted customer engagement but also enhance the chatbot's future performance through insightful analysis.
- **Expanded Wishlist**
  - Expand the existing "wishlist" to encompass not only travel packages but also destination recommendations provided by the chatbot to broader user engagement.
- **Elevated Chatbot UI**
  - Craft a user interface that integrates the chatbot's functionality while aligning with the brand's visual identity. This cohesive design elevates user engagement and reinforces brand recognition.

FUTURE SUGGESTIONS

- **Membership Integration**
  - Incorporate a membership function within the chatbot interface, enabling users to log in or create new accounts.
  - Grant access to account information, including purchase history and chat transcripts, fostering a personalized and seamless experience.
- **Post Matching Interaction**
  - Following successful pairing, inquire whether customers prefer to embark on the journey as per the packaged itinerary or opt for a personalized travel plan.
- **Enriched Travel Expense Prediction**
  - In Scenario 4, offer customers the ability to view estimated travel costs after redeeming "Loyalty Points."
- **Leveraging APIs for Real-Time Precision**
  - In Scenarios 2 and 4, integrate OpenWeather API for precise weather insights and an airfare-related API to offer exact airfare information

APPENDIX

- **Changes to Scenario 1**:
  - In response to feedback, our approach has evolved by expanding destination characteristics from three to five, enhancing recommendation accuracy.
  - Additionally, we've enriched the user interface using HTML, incorporating city introductions and images within recommendations for improved information delivery and engagement.
  - These enhancements align with our goal of providing more accurate and immersive travel suggestions, tailored to individual preferences.

APPENDIX

- **Changes to Scenario 3**:
  - Enhanced K-means Algorithm:
    - Improved accuracy: K-means algorithm upgraded with categorical variable encoding for precise group formations.
    - Personalization boost: Incorporation of language and hobbies variables adds depth to individual preferences analysis.
  - Integrated Destination Recommendation:
    - Travel ideas support: Seamless integration of a destination recommendation system assists indecisive users.
    - Informed choices: Provides tailored suggestions aligning with travelers' interests for enhanced decision-making.
  - Group Member Information Visualization:
    - Transparency enhancement: New visualization feature displays shared interest distribution among group members.
    - Connection facilitation: Offers a clear snapshot of fellow travelers, fostering collaboration and camaraderie.

APPENDIX

- **Changes to Scenario 4**:
  - Strategic Airfare Insights:
    - New comment feature provides timely advice for airfare bookings based on projected price trends.
    - Empowers users to optimize booking decisions and secure cost-effective deals.
  - Enhanced ARIMA Modeling:
    - ARIMA model refinements enhance predictive accuracy for more precise expense projections.
    - Improved trend capturing ensures robust forecasts, contributing to informed planning.
  - Upgraded Data Source and GUI:
    - Integration of upgraded data source increases prediction reliability through enriched data.
    - Refined Graphical User Interface (GUI) ensures an intuitive and seamless user experience.

REFERENCES

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with Python. In Proceedings of the 9th Python in Science Conference (pp. 57-61).
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90-95.
- Amir, H. H. (2020). Wordcloud: A Python package for text mining and word cloud generation. Retrieved from https://github.com/amueller/word_cloud
- Pérez, F., & Granger, B. E. (2007). IPython: A system for interactive scientific computing. Computing in Science & Engineering, 9(3), 21-29.