Restaurant Feedback Analysis

Part 1: Web Scraping [30 Marks]

You are provided with the link for a website from which you have to choose one restaurant. The minimum requirement is for the restaurant to have >1500 reviews. Your job is to scrape reviews about your chosen restaurant from the website using BeautifulSoap and Selenium (handle pagination to ensure all reviews are scraped) and store that data in a meaningful way (for this you'll have to extract additional information such as restaurant name and associated metadata (e.g., dates, ratings)).

Link: Open Table

Requirements

1. Scrape reviews of your restaurant from the target website (OpenTable).
2. Extract the following information:
   - Restaurant name.
   - Review content.
   - Ratings.
   - Date of the review.
3. Save the extracted data in a structured format (e.g., JSON or CSV).

Part 2: Prompt Engineering [20 Marks]

You're provided with skeleton code for loading an LLM model and providing it with data and prompts to generate responses. Use this to separate the food and staff information from the reviews. Vague or poorly constructed prompts can result in outputs that stray from the intended context or requirements. These are termed hallucinations and you must avoid them. Additionally, to improve security of your software, you must avoid saving personal information of the clients even if provided by themselves in the reviews.

Requirements

1. Use the given skeleton code to analyze each review using prompt engineering.
2. Extract and categorize:
   - Comments about food quality.
   - Comments about staff/service.
3. Ensure no hallucinations occur and no data irrelevant to the category is added (e.g. no staff related data should be present in the food information).

4.   Exclude personal information (PI) of reviewers.
5.   Save this data as a .json file.

Part 3: GUI Development [25 Marks]

Build a dashboard using any desired framework (such as streamlit or FLASK) where users can search for food and staff comments for a specific restaurant.

Requirements

1.   Display all reviews in a format similar to the original website.
2.   Highlight the part of each review related to food and staff/service with different colors.
3.   Present the data in a clear, user-friendly manner.

Part 4: Competitor Analysis [25 Marks]

Compare ratings of a selected restaurant with a competitor over time. Do this by extending the scraping functionality to collect ratings over time for selected competitors. Ensure data is in a format suitable for visualization (e.g., date, rating). Visualize the comparison using a graph.

Requirements

1.   Enable users to select a competitor restaurant from opentable by inserting a link on the dashboard.
2.   Scrape ratings data by date for both restaurants.
3.   Plot a time-series graph showing rating trends for both.

Bonus: [10 Marks]
Based on the competitor analysis, generate comparison graphs and a final overall review comparing both restaurants.