“Trip Pulse: Empowering Safe Journeys with Personalized Travel Safety Alerts”

Parichay Madnani (MT23054) Dhananjay Sharma (MT23033) Ummeed Sinha (MT23101)
Megha(MT23125)
Ritik (MT23076) Anupma Pandey(MT23019)

1. Problem Statement

In the realm of modern travel, a significant challenge persists: the absence of easily accessible and comprehensive travel safety information for individuals embarking on journeys. Navigating the complexities of travel planning and execution is fraught with uncertainty, as travelers grapple with the lack of timely updates on potential risks stemming from unpredictable weather conditions, unforeseen events, and limited access to up-to-date news and alerts pertinent to their destinations. This deficiency in real-time information not only jeopardizes the safety of travelers but also subjects them to inconvenience and potentially hazardous situations. Furthermore, the oversight in recognizing the significance of enriching users' travel experiences by highlighting opportunities for exploration and enjoyment exacerbates this issue. To address this pressing need, the development of an application is imperative. This application must serve as a centralized platform, consolidating and disseminating timely updates to empower users with the knowledge necessary to make informed decisions and take proactive measures to safeguard their safety, enhance their enjoyment, and promote their overall well-being throughout their travels. By bridging the gap between travelers and pertinent safety information, this solution seeks to revolutionize the travel experience, ensuring that every journey is not only enriching but also secure and memorable.

2. Motivation

Existing solutions in the realm of travel and tourism predominantly function as intermediary platforms, primarily aimed at fulfilling user needs and generating commissions through transactions facilitated within the app ecosystem. However, this profit-driven approach often sidelines the paramount concern of passenger safety, as the primary objective remains facilitating user travel and bookings through the application.

Another significant issue pertains to the lack of real-time updates regarding destinations. Frequently, existing apps overlook current events transpiring at destinations, thereby failing to provide timely advice to tourists. Instead, they rely solely on historical data, seasonal trends, and past tourist patterns, disregarding unforeseen events such as natural disasters, civil unrest, and local regulations implemented to safeguard travelers' experiences from being disrupted at the eleventh hour.

Furthermore, a notable challenge arises from the reliance on digital news platforms as the primary source of information distribution.

Certain tourist regions in India, for instance, still heavily rely on print media for disseminating information. Consequently, instances like a sudden surge in insurgency in a village in Nagaland may go unreported in digital media channels, resulting in inaccuracies in recommendation scores and potentially exposing tourists to unforeseen troubles. Addressing this issue becomes imperative in subsequent stages of the project to ensure comprehensive and accurate information dissemination.

Moreover, prevailing pre-trained models lack specialization in the news domain, rendering them inadequate in processing the semantics and contexts of news articles. Most of these models are trained on tweet data, thereby failing to capture the intricacies and nuances inherent in news content. This limitation underscores the necessity for tailored solutions capable of comprehensively analyzing news data to facilitate informed decision-making for travelers.

3. Literature Review

The role of various technologies in enhancing travel safety and user experience has garnered significant attention in recent research. This section provides an overview of key studies focusing on social media, recommendation algorithms, government initiatives, and sentiment analysis within the context of travel safety and planning.

3.1 The Role of Social Media in Health Safety Evaluation of a Tourism Destination throughout the Travel Planning Process [1]

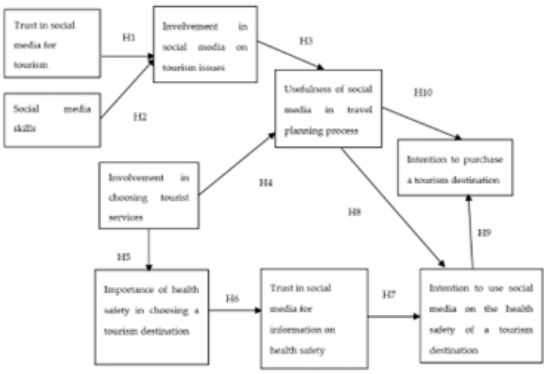

Figure 1. The conceptual model.

This study emphasizes the reliance of travelers on social media platforms for real-time information, reviews, and recommendations concerning health safety measures, particularly in light of the COVID-19 pandemic. By leveraging user-generated content and official announcements, travelers can make informed decisions to safeguard their well-being during travel. However, limitations such as the credibility of information, temporal relevance, and sample bias underscore the need for supplementing social media data with other sources to ensure comprehensive safety evaluation.

3.2 Application of Content-Based Recommendation Algorithms on Mobile Travel Applications [2]

Exploring the application of recommendation systems in the tourism sector, this research highlights the significance of content-based recommendation algorithms in enhancing user experience and destination planning. Despite their effectiveness in providing personalized recommendations, challenges such as dynamic user preferences,

recommendations based on local trends for new users, and computationally expensive processing need to be addressed for optimal performance. Nonetheless, effective recommendation systems play a crucial role in facilitating destination planning for travelers.

3.3 Smart Traveler Enrollment Program (STEP) [3]

The Smart Traveler Enrollment Program (STEP) offered by the U.S. Department of State serves as a valuable resource for travelers by providing up-to-date safety and security information. By subscribing to the program, users receive notifications on relevant travel advisories, enabling them to stay informed about potential risks and hazards at their destinations. Despite its benefits, limitations such as subjective travel advisories and reliance on government updates highlight the need for additional features such as real-time tracking and personalized safety recommendations.

3.4 Rethinking Sentiment Analysis in the News [4]

Traditionally, sentiment analysis has primarily focused on subjective text types with clearly defined targets. This study rethinks sentiment analysis within the news domain, emphasizing the importance of defining targets and separating positive and negative sentiment from news content. However, challenges persist in incorporating semantic and contextual meaning into sentiment analysis models, necessitating further advancements to enhance their efficacy.

In conclusion, while various technologies and initiatives hold promise in enhancing travel safety and planning, addressing their respective limitations is crucial for their effective implementation. By supplementing social media data with other sources, optimizing recommendation algorithms, enhancing government initiatives such as STEP, and advancing sentiment analysis techniques, researchers can contribute to the development of comprehensive solutions that empower travelers to make informed decisions and ensure their safety and well-being during travel.

4. Novelty

Trip Pulse introduces an approach to travel planning by incorporating sentiment analysis of news articles from the traveler's perspective. Unlike traditional methods, Trip Pulse harnesses the power of web-based news articles to gauge traveler sentiments regarding various destinations, revolutionizing trip planning for users.

With a focus on prioritizing the safety perspective of destinations, Trip Pulse goes beyond surface-level recommendations by evaluating crucial safety factors including crime rates, political stability, and health risks. By providing users with comprehensive insights into the safety landscape of destinations, Trip Pulse empowers travelers to make informed decisions about where to go based on their individual safety preferences.

Moreover, Trip Pulse offers personalized insights tailored to each user's unique

preferences, such as interests, budget constraints, and travel history. By leveraging advanced recommendation algorithms, Trip Pulse enhances the overall travel experience by guiding users through every step of the trip planning process, ensuring that their journeys are not only enjoyable but also safe and fulfilling.

5. Methodology

5.1 Data Collection

Our data collection process integrates various APIs and web scraping techniques to gather comprehensive information crucial for our application's functionality. For real-time weather updates, we leverage OpenWeather API, employing Selenium for dynamic web scraping to ensure accuracy and reliability. Additionally, we utilize the OpenCage API to retrieve the precise latitude and longitude of a given destination, facilitating location-based weather analysis. Furthermore, historical weather data spanning the last decade enriches our predictive capabilities, enabling us to assess the suitability of travel conditions for specific destinations. To complement weather insights, we harness SERPAPI to access up-to-date flight and hotel prices, ensuring users have access to comprehensive travel information. Through these meticulously crafted data collection methods, our application delivers informed recommendations tailored to users' travel needs.

5.2 Sentiment Analysis and Recommendation Score Calculation

Sentiment analysis plays a pivotal role in shaping the recommendation score for destinations within the Trip Pulse application. This section talks about the methodology employed for sentiment analysis, specifically tailored to reflect the traveler's perspective on the news content retrieved through scraping. By analyzing the sentiment of news articles from the traveler's viewpoint, we capture nuanced insights into how various destinations are perceived and portrayed in the media. This approach ensures that the recommendation score is not solely based on factual information but also encompasses the subjective experiences and sentiments of potential travelers. The sentiment analysis methodology is then integrated into the calculation of the recommendation score, enriching it with qualitative data and enhancing the relevance of the recommendations provided to users.

5.2.1 Sentiment Analysis Formula

The sentiment analysis process employs a pre-trained model, specifically the 'robertuito-sentiment-analysis' model from Hugging Face's PySentimiento library, to extract sentiment scores from a collection of news headlines. These scores encompass negative, positive, and neutral sentiments, providing a holistic view of the sentiment conveyed within the news corpus. To maintain consistency and comparability across datasets, the scores are normalized to ensure they sum 1up to 1, representing proportions within the headlines. This

normalization procedure guarantees fair assessment and allows each sentiment category to contribute proportionally to the overall sentiment analysis.

5.2.2 Weighting Scheme

Following sentiment analysis, a weighting scheme is applied to assign importance to each sentiment category. Negative weight (WN), positive weight (WP), and neutral weight (WU) coefficients are utilized to reflect the relative significance of negative, positive, and neutral sentiments in influencing the recommendation score.

5.2.3 Temperature Range Influence

In addition to sentiment analysis, the recommendation score is influenced by the user's preferred temperature range for weather suitability. Users specify their preferred temperature range (Tmin to Tmax) to guide the recommendation process effectively.

5.2.4 Recommendation Score Calculation

The recommendation score is calculated using a formula that incorporates the normalized sentiment scores, weighting coefficients, and temperature range influence. The formula is expressed as follows:

RecommendationScore
= (weights[0] * weather_score + weights[1] * sentiment_score + weights[2] * transport_score + weights[3] * hotel_score)

weights = [0.25, 0.35, 0.2, 0.2]

5.3 Database

Our data collection process integrates various APIs and web scraping techniques to gather comprehensive information crucial for our application's functionality. For real-time weather updates, we leverage OpenWeather API, employing Selenium for dynamic web scraping to ensure accuracy and reliability. Additionally, we utilize the OpenCage API to retrieve the precise latitude and longitude of a given destination, facilitating location-based weather analysis. Furthermore, historical weather data spanning the last decade enriches our predictive capabilities, enabling us to assess the suitability of travel conditions for specific destinations. To complement weather insights, we harness SERPAPI to access up-to-date flight and hotel prices, ensuring users have access to comprehensive travel information. Through these meticulously crafted data collection methods, our application delivers informed recommendations tailored to users' travel needs.

6. Code Snippet

get_airport_codes(city_from, city_to):
This function takes two parameters city_from and city_to, representing the departure and arrival cities.
It returns a dictionary containing the corresponding airport codes for the provided cities.
It first normalizes the input by converting city names to lowercase.

Then, it looks up the airport codes in a predefined dictionary and returns them.

fetch_flight_data(departure_id, arrival_id):
This function fetches flight data from a third-party API using the requests library.

It takes departure_id and arrival_id as parameters, representing the departure and arrival airport codes.

It constructs a request URL with specific parameters such as engine, currency, dates, and API key.

It sends a GET request to the API endpoint and returns the JSON response containing flight data.

clean_price(price_str):
This function cleans up a given price string.
It removes any character that is not a digit or comma from the input string.
It returns the cleaned price string.
save_flights_to_csv(flight_data):

This function saves flight data to a CSV file.
It takes flight_data as input, which is a JSON object containing flight information.
It extracts relevant flight details such as airline and price from the JSON response.
It cleans up the price strings using the clean_price function.

It then saves the extracted data to a CSV file named "flight_prices.csv" using pandas DataFrame.

7. Evaluation

7.1 Evaluation Metrics of ML Model

The performance of our machine learning (ML) model, based on the pre-trained robertuito-sentiment-analysismodel provided PySentimiento by Hugging Face, is assessed using two key metrics: F1-score, Precision, Recall, Accuracy. This evaluation was done against manually annotated news. These metrics offer valuable insights into the model's ability to accurately classify news sentiments from the perspective of tourists.

7.2 Evaluating Our Scrapped News Articles

In our endeavor to gather relevant news articles for sentiment analysis, we faced challenges such as retrieving non-relevant articles, such as those pertaining to IPL or political news, which do not directly impact tourism. To address this, we refined our web scraping queries to target only articles pertinent to tourism in specific cities, such as **“**query + " AND [travel + OR + flights + OR + tourist + OR + festival + OR + events + OR + exhibition + OR + carnival + OR + fair**”**. The effectiveness of our web scraping efforts is evaluated using the precision metric. Precision measures the proportion of relevant articles retrieved compared to all articles retrieved, providing valuable insights into the relevance and accuracy of the scraped news articles for sentiment analysis. The evaluation of the recommendation score in our travel application has been conducted through human assessment. Given the novelty of our approach and the absence of established benchmarks for comparison,

human evaluation was deemed the most appropriate method to gauge the effectiveness and relevance of our recommendations. A group of people who have already visited a few destinations recently were tasked with assessing the recommendations provided by our application for evaluation. Their qualitative feedback and subjective judgments served as valuable insights into the perceived usefulness, accuracy, and novelty of our recommendation system. While quantitative metrics may be lacking due to the unique nature of our approach, the human evaluation process offers invaluable qualitative feedback that informs further refinement and enhancement of our recommendation algorithms.

References

[1] Țuclea C-E, Vrânceanu D-M, Năstase C-E. The Role of Social Media in Health Safety Evaluation of a Tourism
Destination throughout the Travel Planning Process. Sustainability. 2020; 12(16):6661.
https://doi.org/10.3390/su12166661

[2] H. Zacarias, G. Cangondo, L. Souza-Pereira, N. M. Garcia, B. Silva and N. Pombo, "Application of Content-Base Recommendation Algorithms on Mobile Travel Applications," 2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC), Jeddah, Saudi Arabia, 2023, pp. 1-5.
https://ieeexplore.ieee.org/document/10085680

[3] Overseas Security Advisory Council (OSAC), "Travel Advisories".
URL:
https://www.osac.gov/Content/Browse/Report?subContentTypes=Alerts%2CTravel%20Advisories

[4] Balahur, Alexandra & Steinberger, Ralf. (2009). Rethinking Sentiment Analysis in the News: from Theory to Practice and back.
https://www.researchgate.net/publication/229052388_Rethinking_Sentiment_Analysis_in_the_News_from_Theory_to_Practice_and_back