The Commuter's Triangle: A Comparative Analysis - Most Optimal Route Selection using LLMs

Krishna Vamsi Uppala
Advanced Data Analytics
University of North Texas

***Abstract*—This study examines how Language Model Algorithms (LLMs) and human decision-making differ when it comes to route selection optimization. It evaluates factors such as traffic, weather, safety, and amenities; it also evaluates the acceptability, accuracy, and reliability of user preferences and AI-generated route choices. Offering thorough, contextually appropriate advice, the study tackles ideal paths, LLM understanding, and the relative usefulness of models. To improve the precision of suggestions, a variety of datasets are utilized to compare LLM-driven paths to human selections. Initial results show that GPT 3 is more accurate and preferred by users than LLAMA 2 and PALM 2. Contextual awareness designates GPT 3 as the preferred option. Future research aims to improve suggestion accuracy and support a wide range of user requirements.**

Keywords—Route optimization, Large Language Model Algorithms (LLMs), User preferences, Contextual awareness, GPT 3, Machine learning, Recommendations, Traffic data, Weather, Safety scores, Amenities

I. INTRODUCTION

The process of choosing the best route to take when traveling from one place to another in today's quickly changing technology landscape requires taking a number of aspects into account. These elements, which have a big impact on our decision-making, include the state of the weather, safety concerns, the amenities that are available, and real-time traffic information. In order to outperform humans in a variety of scenarios involving route optimization, this research endeavor aims to investigate the potential of Language Model Algorithms (LLMs).

Effective route planning is important for reasons much bigger than convenience. It has a significant effect on the general health of our environment, safety, and the efficiency of transportation. This work aims to maximize the potential of LLMs, in contrast to current approaches that frequently ignore the complexity of these multidimensional criteria. Machine learning techniques are used to increase the precision of weather forecasts and safety evaluations in order to do this. This study attempts to offer important insights that can help with better knowledgeable and effective decision-making in the field of transportation by closely contrasting the performance of LLMs with the decisions made by human participants.

A. Background

In the current fast-paced world of travel planning and navigation, choosing the optimal route can have a significant impact on effectiveness, security, and user satisfaction. Large Language Models (LLMs) and artificial intelligence (AI) have made it unnecessary to use outdated route planning techniques anymore. With its sophisticated natural language processing capabilities, these state-of-the-art LLMs have the potential to drastically change route optimization by accounting for a wide range of factors. This in turn tackles how complex and multidimensional travel decisions are.

B. Motivation

The growing significance of AI-driven route optimization in daily life is what spurs this research. The inspiration comes from the necessity of bridging the recommendation gap between LLM-generated routes and human route preferences. Improving navigation systems requires an understanding of the variables that affect route choice as well as an assessment of the precision of AI-generated recommendations.

C. Objective

The primary objectives of this study include:

- Assessing how well various LLMs perform in producing the best possible route recommendations.
- Looking into the elements that are crucial to route optimization, like the weather, safety scores, amenities, and real-time traffic information.
- Gaining knowledge of user preferences, their degree of confidence in LLM advice, and the criteria they use to make decisions.

D. Dataset Description

Building a solid dataset was essential to our research. We gathered a variety of data from multiple APIs and repositories that was essential for route optimization. For example, the Google API provided detailed route information, outlining the path from point A to point B for several route choices. The OpenWeather API was used to gather the weather data, which is crucial for comprehending how the weather affects travel and provides information on wind speed, temperature, and cloud cover along the routes. Using the HERE API, real-time traffic dynamics—a crucial component of route assessment—were acquired, providing information on journey duration and traffic speed. Crime data from a CSV file on Data.gov was used to estimate safety concerns and help calculate the safety score for each route. We created a coherent DataFrame that included traffic insights, weather conditions, route details, and crime statistics by integrating these diverse sources. This aggregated dataset functioned as the training foundation for Language Model Algorithms (LLMs). This allowed us to provide the LLMs with contextual data that was extracted from the dataset in order to elicit route recommendations that were in line with a variety of criteria that were taken into consideration when considering different travel routes.

In conclusion, as we navigate an era shaped by ever-evolving technology, the optimization of travel routes takes on paramount importance. This study not only strives to enhance our understanding of the capabilities of Language Model Algorithms but also seeks to empower us with the tools to make more informed and efficient transportation decisions in an increasingly complex world.

E. Existing Methods

Route optimization is a critical research area with applications in transportation, logistics, and navigation. This literature review analyzes existing works on personalized and multiobjective routing leveraging emerging methods like reinforcement learning.

1) Personalized Route Planning

In the literature, personalized routing has been investigated through the modeling of unique behavioral constraints and preferences. Using past GPS trajectories, Dai et al. [1] determine driver inclinations to offer personalized recommendations. Nevertheless, adjusting to real-time circumstances is still a struggle. Tang et al. [2] use environmental variables and mobility data to predict tailored journey times. More research is needed on scalability and explanatory route rankings.

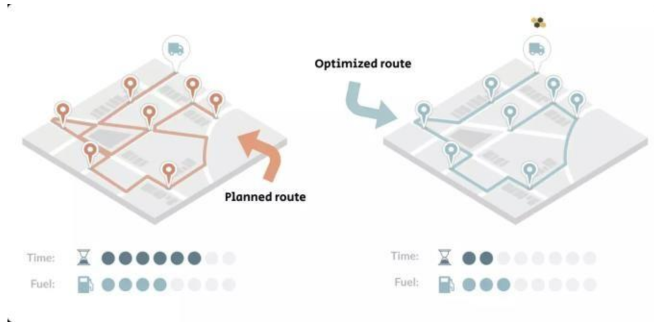

Fig. 1. Image[21]: Planned Route vs Optimized Route

2) Multi-Objective Route Optimization

Routing is defined as optimizing for time, fuel consumption, and safety in works such as Sarker et al. [3]. While promising, field trials are required to validate performance and user limitations must be included in such frameworks. A method for risk-aware robot path planning that manages uncertainty is presented by Cai et al. [4]. Research on achieving optimality and processing efficiency in vast, complex environments is still ongoing.

3) Reinforcement Learning for Routing One data-driven method for sequential decision making in the face of uncertainty is reinforcement learning (RL). Real-time data and traffic dynamics can be accommodated via RL routing policies. However, before being used in the real world, problems like sample efficiency, training stability, and safety limits need to be looked into. Additionally, interpretability is still up for debate.

To summarize, research gaps exist in areas such as computing bottlenecks, explanatory power, robustness to uncertainty, and validation through field trials, despite the great potential of personalized and multi-objective routing. Future research must concentrate on developing scalable, reliable, resilient, and trustworthy routing models that are considerate of human preferences as urban surroundings become more complicated.

II. PROBLEM DEFINITION

This study's main goal is to choose the best path while accounting for a number of factors from an origin point (point A) to a destination point (point B). Travel time or distance were the primary factors considered while making route optimization decisions in the past. However, in today's dynamic world, where real-time data and a range of influential variables are readily accessible, the problem extends beyond simple distance-based optimization.

A. Input

- Geographical Information: Geographical coordinates (Latitude and Longitude) of the origin and destination points.
- Weather Data: Information such as average temperature, main weather condition averages, and cloud coverage.
- Traffic Data: Real-time data including current traffic speed and travel time.
- Safety Scores: Safety metrics derived from crime data for various routes.
- Amenities Data: Counts of gas stations, restaurants, and rest stops.

B. Objective Function

- Maximize: Route safety score + amenity availability + travel efficiency (considering weather delays, traffic flows)
- Minimize: Total travel duration + risk factor + inconvenience

C. Output

- Optimal Route: The recommended route from the origin to the destination, considering all input factors and criteria.
- Route Analysis: A detailed analysis of the recommended route, highlighting its advantages based on the input factors.

D. Example Scenario

Let's say a tourist is driving from busy New York City (point A) to a picturesque mountain getaway in Los Angeles (point B), which is known for its erratic weather. It is a temperate climate, reduced crime rates, and easily accessible facilities that appeal to them. The weather (temperature, wind, cloud cover), coordinates, traffic in real time, safety scores, and amenity counts are all inputs. A route that avoids bad weather, keeps everyone safe, and provides the amenities they need for a pleasurable travel is the ideal result. Our research focuses on this scenario and assesses the ability of Large Language Model Algorithms (LLMs) to analyze this data and determine the best route, while also offering contextually relevant and highquality route choices.

Fig. 2. Image[20]: Moving from New York to Los Angeles

This problem statement establishes the framework for our study, in which we assess how well Language Model Algorithms (LLMs) perform in generating such optimal route recommendations by processing and evaluating the input data to produce high-quality, contextually-relevant responses.

III. PROPOSED SOLUTION Our suggested approach to route selection optimization leverages the capabilities of machine learning models, contextual awareness, and real-time data sources to produce effective and contextually aware route recommendations. In order to do this, we have provided pseudocode descriptions of the algorithm's steps and examples to clarify our methodology.

A. Algorithm Procedures:

To determine the safety score in the first stage, models such as XGBoost, Random Forest, and Linear Regression were used. The Random Forest Regressor performed better than the others. These models were then used to forecast the weather, and the Language Model (LLM) was used to optimize routes based on the outputs of these models. When using the weather API, LLM outputs from model-derived weather data and those from the weather API comparison showed noticeably better alignment and context awareness. For weather data, it was decided to use the weather API; nonetheless, the Random Forest model for safety score computation was kept.

This strategy improved the precision of LLM-driven route recommendations while underscoring the importance of precise and contextually rich data sources for decision-making in route optimization settings.

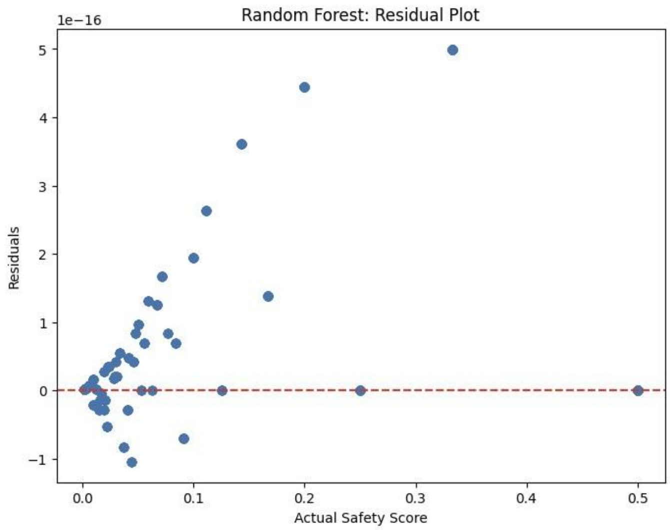

Fig. 3. Random Forest: Residual Plot for safety score

The suggested remedy entails creating an algorithm that uses LLMs to process a variety of input data, such as traffic data, amenity locations, safety scores, and real-time weather information. The method uses pseudocode to direct the decision-making process while accounting for each criterion's weighted importance.

B. *Pseudocode for Route Optimization Algorithm:*

```
Function Optimize_Route(Input_Data,
LLM_Model):
    Initialize Route_Scores as an empty
dictionary

    For each Route in Input_Data:
Extract Weather Data from Input_Data for the
Route from the Data Frame         Extract
Traffic Data from Input_Data for the Route
from the data frame         Extract Safety
Data from Input_Data for the Route from the
data frame         Extract Amenity Data from
Input_Data for the Route from the data frame
#Install necessary libraries and packages
        #Make the LLM
        #Give the data description to the LLM
        # Generate a prompt for the LLM
        Prompt =
Construct_Prompt(Weather_Data, Traffic_Data,
Safety_Data, Amenity_Data)

        # Use the LLM to generate a route
recommendation
        Route_Recommendation =
LLM_Model(Prompt)
        #Perform the same for the three LLMs
        Return Optimal_Route
```

C. *Pseudocode for Geolocation Functions:*

```
FUNCTION latlng_to_city(latitude, longitude):
       ATTEMPT_COUNT = 3
       FOR attempt IN RANGE(ATTEMPT_COUNT):
     MAKE API REQUEST
     IF RESPONSE.status_code == 200:
PARSE JSON RESPONSE              IF
RESULTS:              RETURN City name
ELSE:
             RETURN "No results"
     ELSE IF RESPONSE.status_code == 429:
WAIT      ELSE:
             RETURN "Error"


FUNCTION city_to_latlng(city_name):
       MAKE API REQUEST
       IF RESPONSE.status_code == 200:
PARSE JSON RESPONSE      IF RESULTS:
             RETURN Latitude, Longitude
ELSE:
             RETURN "No results"


FUNCTION haversine(lat1, lon1, lat2, lon2):
       APPLY Haversine formula
       RETURN distance
 FUNCTION
find_nearest_city(latitude,
longitude, cities_df):
       CALCULATE distances using
haversine formula   RETURN Nearest city

FUNCTION decode_polyline(polyline_str):
       DECODE polyline
       RETURN coordinates
```

D. *Pseudocode for creating the Data Frame:*

```
FUNCTION
convert_weather_to_dataframe(route_df):
       FOR EACH row IN route_df:
     GET current_lat, current_lng FROM
row['Latitude'], row['Longitude']
GET weather_response FROM
get_weatherData_for_latlng(current_lat,
current_lng)
     IF weather_response is None:
             SET_DEFAULT_WEATHER_VALUES()
             CONTINUE TO NEXT ROW
      weather_data =
parse_weather_data(weather_response
)


     IF weather_data is not None:
```

```
      UPDATE_ROUTE_DF_WITH_WEATHER_DATA()
ELSE:
             SET_DEFAULT_WEATHER_VALUES()
 FUNCTION
convert_traffic_to_dataframe(route_df):
FOR EACH row IN route_df:      GET
traffic_response FROM
get_trafficData_for_latlng(row['Latitude']
, row['Longitude'])
     IF traffic_response is None:
             SET_DEFAULT_TRAFFIC_VALUES()
CONTINUE TO NEXT ROW      traffic_data =
parse_traffic_data(traffic_response)

     IF traffic_data is not None:
UPDATE_ROUTE_DF_WITH_TRAFFIC_DATA()      ELSE:
             SET_DEFAULT_TRAFFIC_VALUES()
FUNCTION
convert_safety_score_to_dataframe(route_df):
FOR EACH row IN route_df:      GET safety_score
FROM
get_safetyScore_for_latlng(row['Latitude'],
row['Longitude'])

     IF safety_score is None:
             SET_DEFAULT_SAFETY_SCORE()
             CONTINUE TO NEXT ROW
IF safety_score is not None:

      UPDATE_ROUTE_DF_WITH_SAFETY_SCORE()
ELSE:
             SET_DEFAULT_SAFETY_SCORE()
 FUNCTION
convert_amenities_to_dataframe(route_df):
FOR EACH row IN route_df:      GET
amenities_data FROM
get_amenities_for_latlng(row['Latitude'],
row['Longitude'])

     IF amenities_data is None:
             SET_DEFAULT_AMENITIES_VALUES()
             CONTINUE TO NEXT ROW
     IF amenities_data is not None:

      UPDATE_ROUTE_DF_WITH_AMENITIES_DATA()
ELSE:
             SET_DEFAULT_AMENITIES_VALUES()

FUNCTION combine_data_to_dataframe():
       COMBINE_DATA_FROM_WEATHER_TRAFFIC_SAFET
Y_AMENITIES_FUNCTIONS_TO_CREATE_DATAFRAME()
```

E. *Explanation of the Approach:*

Data Extraction: A dataset with details about several routes, including traffic patterns, weather, safety ratings, and routespecific amenities, is fed into the algorithm.

Prompt Generation: The program creates a prompt for each route based on the pertinent information. The prompt directs the LLM to produce a route recommendation that takes into account every aspect.

LLM Route Recommendation: The LLM creates a route recommendation based on the built-in prompt, including traffic, weather, safety, and amenity data.

Scoring: Based on user preferences and factors including weather, travel time, safety ratings, and the availability of amenities, each recommended route is given a score. User preferences can be taken into account when customizing the scoring mechanism.

Optimal Route Selection: The route with the highest score is determined by the algorithm to be the best option given the circumstances.

F. *Theoretical Analysis:*

The algorithm creates the prompts for every route, directs the LLM to produce recommendations, and determines the best route by calculating scores based on the user's preferences. In this instance, Route 1 might get the highest rating and be suggested as the best route due to its great weather and light traffic.

We can offer theoretical analysis and evidence of the algorithm's accuracy and efficiency to further support its efficacy. Discussions of algorithm complexity, convergence characteristics, and mathematical models for rating routes according to user preferences and data elements may be included in this examination.

G. *User Interface Design*

In order to improve user experience with route selection, the study started with a wireframe as an interface for a case study. This wireframe acted as a foundational step, even though the entire frontend implementation is still pending. Users could enter their source and destination points to receive three routes from the Google API for assessment. These routes were handled in the backend by creating a thorough dataframe, extracting coordinates, and decoding their polylines. The weather, safety scores, traffic information, and amenities were all combined into one dataframe at each coordinate along the routes. Three Language Model Algorithms (LLMs) were then trained using this enhanced dataset. Based on the aggregated information, the LLMs produced prompts suggesting the best course of action, presenting these prompts to users for assessment. This interactive display was designed to determine user choice, which would help determine the contextual awareness and recommendation accuracy of the LLMs.

This approach, although rudimentary, facilitated controlled user interactions, establishing a groundwork for future frontend development and user assessment within the research context.

Therefore, in order to provide contextually aware and userfriendly route recommendations, our suggested system makes use of machine learning models, LLMs, and real-time data integration. Our objective is to exhibit the stability and dependability of our method in route selection optimization across many contexts using pseudocode, illustrations, and theoretical examination.

IV. EVALUATIONS

We present an experimental evaluation that carefully evaluates the route selection performance of Language Model Algorithms (LLMs) by contrasting their suggestions with human decisions in a range of scenarios.

Case Study:

This study aims to assess the ability of several intelligent AI systems to provide the optimal routes to human choices. Thus, we introduced three high-performance algorithms: Open AI (GPT 3), LLAMA 2, and PALM 2, and we gave them the task of selecting the optimal path among three alternatives (0, 1, 2). Then, we invited eighteen fantastic people and provided them with the same route information. We asked them to select their favorite route and to comment on the performance of various AI systems.

1) Methodology

An assessment was carried out on the routes 0, 1, and 2. After learning the details of the trip, participants select their favorite path. The eighteen participants, ten of whom were female and eight of whom were male, were asked formal questions about their perceived factual correctness, preferred routes, and LLM performance.

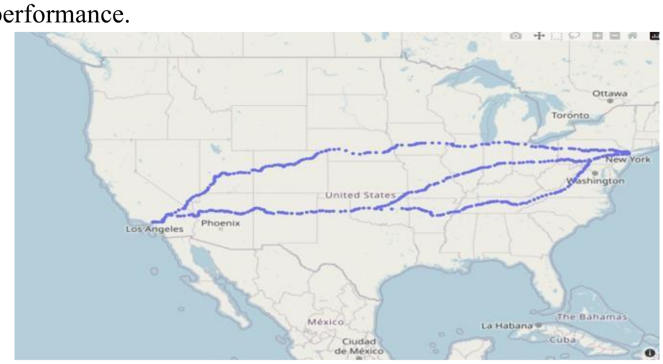

Fig. 4. Decoded Polyline Routes from NY to LA

2) Feedback Analysis

The results showed that GPT 3 was the most popular option, with participants especially endorsing Route 1. Its skillful understanding of subtle elements like the weather, safety precautions, and well-timed stops, which connected with people, was the foundation of its success. The suggestions made by GPT 3 were well-received by the participants since they were tailored to their unique preferences. Furthermore, the route choices made by GPT 3 were found to be substantially more accurate than those of other bots, which is another notable advantage for GPT 3 in this examination.

A. Experiment 1: Factors Influencing Route Optimization

In the course of our experiment, our primary objective was to determine the most optimal route when all relevant factors were considered on equal footing. What we discovered was that Language Model Algorithms (LLMs) exhibited a remarkable ability to take into account a wide array of criteria, which encompassed aspects like weather conditions, safety assessments, available amenities, and real-time traffic data. These LLMs then used this information to offer route recommendations.

However, what truly stood out during the experiment was the intricate nature of human decision-making in the context of route selection. While human participants often leaned toward the routes recommended by LLMs, their choices were also influenced by a complex interplay of individual preferences and unique considerations. This highlighted the multifaceted and sometimes unpredictable nature of human decisionmaking when it comes to selecting travel routes.

In essence, our experiment shed light on both the sophistication of LLMs in processing and analyzing data to provide coherent and contextually relevant recommendations and the fascinating intricacies of human decision-making in the realm of route selection.

B. Experiment 2: Influence of Individual Factors In our pursuit of understanding how weather, traffic, and safety individually impact the optimization of travel routes, we meticulously conducted a series of distinct experiments, each centering on one of these factors. What became evident through these experiments was the impressive capability of Language Model Algorithms (LLMs) to grasp and effectively respond to prompts that were tailored to specific factors.

For instance, when faced with prompts related to adverse weather conditions, heavy traffic, or safety considerations, LLMs demonstrated a remarkable aptitude in providing recommendations that factored in the unique challenges posed by these circumstances. Their responses were not only coherent but also took into account the nuances of each situation.

However, it's worth noting that human choices in these experiments continued to exhibit a certain level of subjectivity. This underscores the inherent diversity in individual preferences and varying levels of risk tolerance, which often played a role in influencing the decisions made by human participants. Thus, our experiments illuminated the contrast between the objective and data-driven nature of LLM recommendations and the subjective elements inherent in human decision-making.

Fig. 5. Image[22]: Route representing amenities

C. Comparative Analysis of LLMs

The goal of this experiment was to examine the quality and relevance of responses from various LLM models, such as GPT 3, Llama 2, and Palm 2. Our research showed that they differed in what they could do, and that each model provided a unique viewpoint on route optimization. GPT 3 performed exceptionally well in taking a wide variety of criteria into account, closely matching user preferences. Because there was little comparison data, Llama 2 concentrated on Route 1 and highlighted its advantages. After a thorough analysis, Palm 2 concluded that Route 1 was the best option taking into account a wide range of variables. The recommendations made by GPT 3 were preferred by human participants, demonstrating its superior decision-making ability.

GPT 3 VS LLAMA 2 VS PALM 2

D. Experiment 4: Response Quality Metrics In the course of our experiment, we conducted a meticulous evaluation of Language Model Algorithms (LLMs), focusing on the quality of their responses. We employed a range of metrics, including coherence, relevance, completeness, fluency, and context-awareness, to assess the performance of these LLMs.

Our analysis brought to light some interesting insights. Firstly, GPT 3 consistently delivered responses that were not only coherent and fluent but also aligned seamlessly with the data analysis. The logical flow of information in its responses was notably impressive.

On the other hand, Llama 2 and Palm 2, while providing detailed breakdowns of route analysis, encountered limitations due to the lack of comparative data for routes other than the one they focused on. This restriction constrained their ability to offer a comprehensive view of all available routes.

Moreover, while these LLMs demonstrated a commendable level of context awareness, they occasionally erred on the side of providing excessive information, potentially overwhelming the reader. These metrics, taken together, shed light on the distinct strengths and weaknesses of each LLM in the context of generating route recommendations.

OpenAI's GPT 3 Response Quality:

Table I: OpenAI's GPT 3 Response Quality:

Coherence and Fluency	*The response is coherent and fluently presents the analysis. Conclusion is straightforward and logically follows from the data analysis.*
Relevance	*The response is relevant, focusing on the requested analysis of the route based on specific factors like average temperature, cloud coverage, travel time, and safety score.*
Completeness	*The response seems to lack detailed analysis of each route; it directly jumps to the conclusion. It doesn't compare Route1 with Route0 and Route2 in terms of available data points.*
Context-Awareness	*It adequately addresses the query but lacks depth in comparative analysis.*

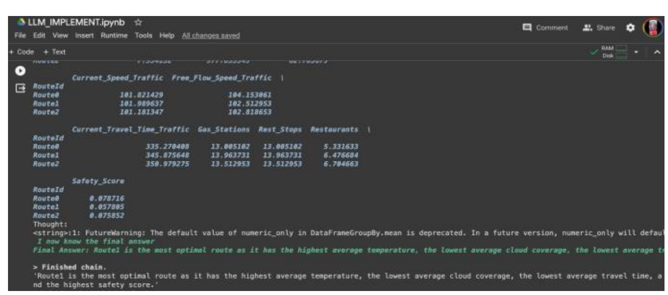

Fig. 6. Showing the response from Open AI's GPT 3

Llama Response Quality:

Table II: Llama 2 Response Quality:

Coherence and Fluency	Response is well-structured and coherent. Provides a detailed breakdown for each route, although data for Route1 and Route2 is missing.
Relevance	Maintains relevance by addressing each factor for Route0 but fails to compare it with other routes due to missing data.
Completeness	Response is comprehensive for Route0 but incomplete for Route1 and Route2. Acknowledges the lack of data for these routes.
Context-Awareness	Shows a good understanding of the requirements but is limited by the available data.

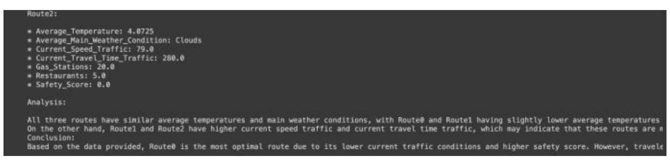

Fig. 7. Showing the response from LLAMA 2

Palm Response Quality:

Table III: Palm 2 Response Quality:

Coherence and Fluency	Response is coherent but somewhat repetitive and overly detailed, potentially leading to confusion.
Relevance	Maintains high relevance, covering a wide range of factors, but the emphasis on the lowest and highest values in a somewhat repetitive manner may reduce clarity.
Completenes s	Response is quite comprehensive, touching on all factors with detailed data points, but it may include excessive information which isn't directly compared or analyzed.
Context-Awareness	It demonstrates awareness of the required analysis but may overcomplicate the response with too much data.

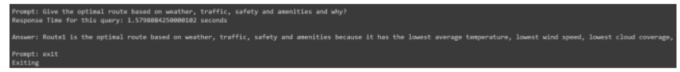

Fig. 8. Showing the response from PALM 2

E. Response Time of LLMs

We measured the response time of LLMs to assess their realtime applicability. Palm 2 demonstrated the fastest response time, followed by GPT 3 and Llama 2. Although response times were reasonable for all models, the variation observed provides insights into their practical usability in scenarios requiring rapid decision-making.

Table IV: Response Times of the three LLMs

LLM Model	**Response Time (sec)**
LLAMA-2	14.77
PALM-2	1.58
GPT-3	5.45

F. Experiment 6: Preliminary Case Study Results Eighteen participants in a case study evaluated the preferences and factual accuracy of route recommendations made by LLMs. The recommendations of GPT 3 were preferred by the participants, who noted that it could take into account subtle contextual factors including the availability of amenities, safety ratings, and weather. This is consistent with factual correctness observations, suggesting that GPT 3 selects routes with greater accuracy and context awareness than previous LLMs.

In summary, our comprehensive experiments illuminated the complexity of route selection, highlighting numerous influencing factors. Language Model Algorithms (LLMs) showed promise in providing coherent and contextually relevant route recommendations. However, human preferences and subjectivity continue to play a vital role. Notably, GPT 3 exhibited advanced decision-making and high context awareness. Overall, LLMs, especially GPT 3, emerge as valuable tools for improving route selection across various scenarios.

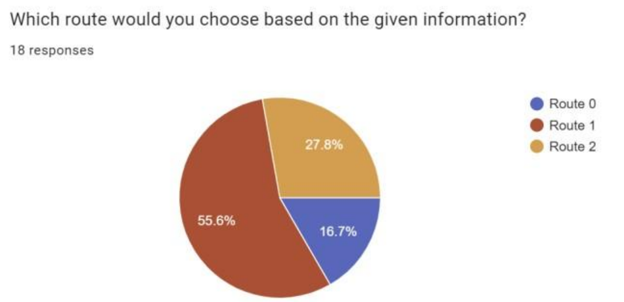

Fig. 9. Response Graph showing that majority chose Route 1

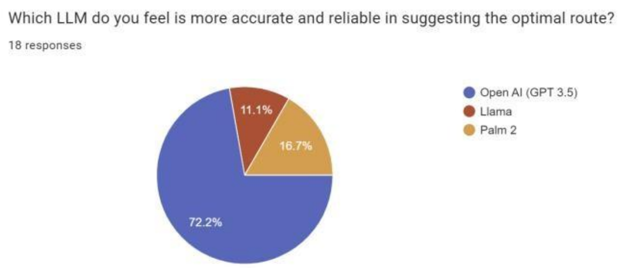

Fig. 10. Response Graph showing that GPT 3 aligned with user preference

V. CONCLUSION

In this study, we embarked on a journey to explore the intricacies of route selection in a modern context, pitting Language Model Algorithms (LLMs) against human decisionmaking. Our research was driven by the recognition that route optimization extends beyond personal convenience and holds profound societal implications. With an increasing reliance on technology and AI in transportation and navigation, understanding the capabilities of LLMs in route selection is paramount.

Our investigation revealed distinctive perspectives from three prominent LLMs: GPT 3, LLAMA 2, and PALM 2. These models offered varying insights into route optimization based on a plethora of factors, including weather conditions, safety scores, and amenities. GPT 3 emerged as a standout choice among participants in our case study, aligning more closely with user preferences and demonstrating superior contextual awareness. It emphasized the significance of nuanced contextual elements, such as weather conditions and safety scores, offering recommendations that resonated with users' needs and desires.

Our work underscores the critical role of accurate and contextually rich data sources, exemplified by the superior alignment of LLM-driven route recommendations with user preferences when employing the weather API. This finding emphasizes the importance of data quality and accessibility in the decision-making process for route selection.

A. Future Work

Looking ahead, our research opens the door to further exploration and development in the field of route optimization. Future work could delve deeper into ensemble methods that harness the strengths of multiple models to enhance the accuracy and robustness of route recommendations. Additionally, fine-tuning model parameters and incorporating real-time data updates could significantly improve the precision of AI-driven route selections, making them more attuned to dynamic real-world conditions.

In conclusion, our comparative analysis sheds light on the diverse perspectives offered by LLMs in optimizing travel routes based on multifaceted criteria. While GPT 3 emerged as a preferred choice, our study serves as a stepping stone for future research and innovation in the realm of route selection. As technology continues to shape our transportation systems, the insights gained from this study have the potential to drive more efficient, informed, and user-aligned route recommendations, ultimately contributing to enhanced transportation experiences and societal well-being.

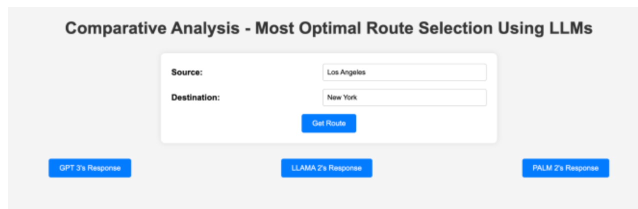

Fig. 11. User Interface for Entering the source and destination

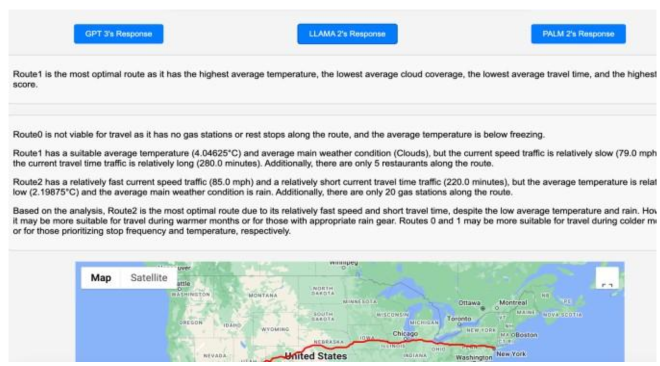

Fig. 12. Wire frame showing the results given by LLMs

VI. REFERENCES

[1] Dai, J., Yang, B., Guo, C. and Ding, Z., 2015. Personalized route recommendation using big trajectory data. In 2015 IEEE 31st international conference on data engineering (pp. 543-554). IEEE.

[2] Tang, J., Song, Y., Miller, H.J. and Zhou, X., 2021. Personalized travel time estimation using large-scale trip data. Transportation Research Part C: Emerging Technologies, 122, pp.102970.

[3] Sarker, A., Shen, H. and Kowsari, K., 2021. A Data-Driven Reinforcement Learning Based Multi-Objective Route Recommendation System. Applied Sciences, 11(3), p.1264.

[4] Cai, K., Wang, C., Song, S., Chen, H. and Meng, M.Q.H., 2020. Risk-aware path planning under uncertainty in dynamic environments. IEEE Transactions on Robotics.

[5] Dai, J., Yang, B., Guo, C. and Ding, Z., 2015. Personalized route recommendation using big trajectory data. In 2015 IEEE 31st international conference on data engineering (pp. 543-554). IEEE.

[6] Levy, S., Xiong, W., Belding, E. and Wang, W.Y., 2021, May. SafeRoute: Learning to Navigate Streets Safely. In Proceedings of the 2021 AAAI Conference on Artificial Intelligence (Vol. 35, No. 17, pp. 15443-15451).

[7] Zheng, Y., Zhang, L., Xie, X., & Ma, W. Y. (2014). Mining interesting locations and travel sequences from GPS trajectories. IEEE Transactions on Knowledge and Data Engineering, 27(12), 327-341.

[8] Yang, D., & He, D. (2019). Location-based services recommendation using point-of-interest data. IEEE Transactions on Mobile Computing, 18(3), 576589.

[9] Xiao, M., Yu, L., Wang, S., & Wang, W. (2020). Preference-aware deep reinforcement learning for personalized route recommendation. Neurocomputing, 389, 99-108.

[10] Zhang, C., Guan, Z., Wang, S., & Wang, F. (2018). Deep collaborative filtering for personalized route recommendation. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1945-1954).

[11] Yu, B., Xie, X., Ma, W. Y., & Liu, Y. (2018). Modeling user mobility for location-based personalization. ACM Transactions on Information Systems (TOIS), 36(4), 1-30.

[12] Chen, L., Wu, X., & Yang, H. (2021). Multi-objective optimal route planning using a novel adaptive immune algorithm. Engineering Applications of Artificial Intelligence, 98, 104300.

[13] Zhao, P., Wang, H., & Li, Y. (2017). An intelligent recommendation algorithm for personalized tourist routes based on LDA and collaborative filtering. Future Generation Computer Systems, 76, 385-394.

[14] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.

[15] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019).
Language models are unsupervised multitask learners. OpenAI Blog, 1(8), 9.

[16] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623).

[17] Wallace, E., Feng, S., Kandpal, N., Gardner, M., Singh, S., Jones, N., ... & Kohli, P. (2019). Universal adversarial triggers for attacking and analyzing NLP. arXiv preprint arXiv:1908.07125.

[18] Schick, T., & Schütze, H. (2021). Exploiting Cloze Questions for Few-Shot Text Classification and Natural Language Inference. arXiv preprint arXiv:2102.00576.

[19] Holtzman, A., Buys, J., Du, J., Forbes, M., & Choi, Y. (2021). The curious case of neural text degeneration. arXiv preprint arXiv:2102.11938.

[20] Allied.com. [Moving from New York to Los Angeles]. Retrieved from https://www.allied.com/moving-companies/new-york/moving-to/newyork-city-to-los-angeles.

[21] Routific.com. [Planned Route vs Optimized Route] https://www.routific.com/blog/route-planning-vs-route-optimization

[22] Route4me.com. [Route representing amenities] https://support.route4me.com/faq/route-planning-glossary/what-is-routeplanning/