Uso di tecniche di LLM per analizzare traiettorie turistiche e per suggerire la prossima attrazione turistica da visitare

Simone Mattioli

Department of Computer Science

University of Verona

July 18, 2025

Abstract

This paper presents a comprehensive study on leveraging Large Language Models (LLMs) for predicting human mobility patterns in urban tourism contexts. We introduce LLM-Mob, a novel framework that utilizes the natural language understanding capabilities of modern LLMs to predict tourists' next Points of Interest (POI) based on their historical visiting patterns. Our approach is evaluated on the VeronaCard dataset, which contains real-world tourism data from Verona, Italy, spanning multiple years. We investigate the impact of different contextual information types on prediction accuracy, including basic POI names, geographical coordinates, and temporal features. The framework demonstrates the potential of LLMs to understand complex spatial-temporal patterns in human mobility without requiring extensive feature engineering or domain-specific architectures. Our implementation replaces traditional OpenAI GPT models with open-source Llama 3.1, making the approach more accessible and cost-effective for research purposes.

1 Introduction

Human mobility prediction has emerged as a crucial research area with applications spanning urban planning, transportation systems, recommendation systems, and public health. Traditional approaches to mobility prediction have relied heavily on statistical models, machine learning algorithms, and deep learning architectures specifically designed for sequential data. However, recent advances in Large Language Models (LLMs) have opened new avenues for understanding and predicting human behavior patterns through natural language processing capabilities.

The fundamental challenge in human mobility prediction lies in capturing the complex interplay between spatial, temporal, and behavioral factors that influence individual movement decisions. Traditional machine learning approaches often require extensive feature engineering and domain-specific knowledge to encode these relationships effectively. In contrast, LLMs possess inherent capabilities for understanding contextual relationships and patterns, which can be leveraged to predict mobility patterns through carefully designed prompts and contextual information.

This paper introduces LLM-Mob, a framework that harnesses the power of Large Language Models for human mobility prediction in tourism contexts. Our approach differs from conventional methods by treating mobility prediction as a natural language understanding task, where historical visiting patterns, geographical information, and temporal context are encoded as textual prompts for LLM inference.

The primary contributions of this work include:

- A novel framework for human mobility prediction using Large Language Models that requires minimal feature engineering

- Comprehensive evaluation of different contextual information types on prediction accuracy

- Implementation using open-source Llama 3.1 models, eliminating dependencies on proprietary APIs

- Extensive experiments on real-world tourism data from the VeronaCard dataset

- Analysis of the impact of geographical proximity and temporal patterns on mobility prediction performance

2   Related Work

Human mobility prediction has been extensively studied across multiple disciplines, with approaches ranging from statistical models to deep learning architectures. Early works focused on Markov-based models that captured transition probabilities between locations based on historical patterns. More recent approaches have leveraged recurrent neural networks (RNNs), Long Short-Term Memory (LSTM) networks, and attention mechanisms to model sequential dependencies in mobility data.

Graph-based approaches have gained prominence due to their ability to capture spatial relationships between locations. These methods typically represent locations as nodes in a graph, with edges encoding spatial proximity or transition probabilities. However, such approaches often require extensive preprocessing and domain-specific feature engineering to achieve satisfactory performance.

The emergence of Large Language Models has introduced new possibilities for mobility prediction. Recent works have explored the use of LLMs for various spatial-temporal tasks, demonstrating their ability to understand geographical relationships and temporal patterns through natural language descriptions. Our work builds upon these foundations by specifically focusing on tourism mobility prediction and investigating the impact of different contextual information types.

3   Methodology

3.1   Problem Formulation

We formulate the human mobility prediction problem as follows: Given a tourist's historical sequence of visited POIs $S = \{p_1, p_2, ..., p_n\}$ ordered by visit time, we aim to predict the next POI p_{n+1} that the tourist is likely to visit. Each POI p_i is characterized by its

name, geographical coordinates (latitude and longitude), and temporal information (visit timestamp).

The prediction task is approached as a natural language generation problem, where the LLM is provided with contextual information about the tourist's behavior and asked to generate the most likely next destinations.

3.2 Dataset Description

Our experiments are conducted on the VeronaCard dataset, which contains real-world tourism data from Verona, Italy. The dataset encompasses visitor logs from 2014 to 2020, capturing tourist interactions with various cultural and historical sites throughout the city. The dataset includes:

- Tourist card usage logs with timestamps and POI identifiers

- A comprehensive catalog of 70 Points of Interest with geographical coordinates

- Visitor patterns spanning multiple years, providing insights into seasonal and temporal variations

The POI catalog includes major tourist attractions such as Arena di Verona, Casa di Giulietta, Basilica di San Zeno, and Castelvecchio Museum, among others. Each POI is associated with precise geographical coordinates, enabling spatial analysis of tourist movement patterns.

3.3 Data Preprocessing and Clustering

Our preprocessing pipeline consists of several stages designed to clean and structure the raw mobility data:

3.3.1 Data Cleaning and Filtering

Raw visitor logs are processed to extract relevant information including visit timestamps, POI names, and tourist card identifiers. We filter out incomplete records and focus on tourists with multiple POI visits to ensure meaningful sequence analysis.

3.3.2 User Clustering

To capture behavioral patterns among tourists, we employ K-means clustering on user-POI interaction matrices. The clustering process involves:

1. Construction of a user-POI matrix where each row represents a tourist and each column represents a POI

2. Standardization of the matrix using StandardScaler to ensure equal weighting of features

3. Application of K-means clustering with $k = 7$ clusters to group tourists with similar visiting patterns

This clustering approach enables the identification of distinct tourist profiles, such as culture-focused visitors, history enthusiasts, or casual tourists, which can inform the prediction process.

4 Prompting Strategies and Experimental Results

The efficacy of Large Language Models (LLMs) in next-POI prediction is fundamentally dependent on the design of context-aware prompting strategies that effectively encode spatio-temporal mobility patterns. Our experimental framework systematically evaluates three distinct prompting paradigms, each progressively incorporating additional contextual dimensions to assess their impact on recommendation accuracy.

4.1 Prompting Strategy Taxonomy

We define three hierarchical prompting strategies, each building upon the previous one to evaluate the incremental contribution of different contextual features:

1. **Baseline Strategy (POI Name Only):** The foundational approach provides the model with a chronologically ordered sequence of previously visited POIs represented solely by their canonical names. This strategy serves as our baseline, focusing purely on sequential patterns without any additional contextual information.

2. **Geospatial-Enhanced Strategy (Name + Geolocation):** Building upon the baseline, this approach enriches each POI with precise geographic coordinates (latitude, longitude). The incorporation of geospatial information enables the model to leverage spatial proximity patterns and geographical clustering effects inherent in tourist mobility behavior.

3. **Comprehensive Context Strategy (Name + Geolocation + Temporal Context):** The most sophisticated approach integrates temporal features alongside spatial information, including precise timestamps, time-of-day categorizations, and day-of-week indicators. This strategy captures the full spatio-temporal complexity of tourist mobility patterns.

4.2 Experimental Protocol

Our evaluation employs a systematic approach utilizing the VeronaCard dataset, which provides comprehensive tourist mobility traces across multiple years (2014-2020). The experimental protocol incorporates the following key components:

- **Clustering-based User Segmentation:** Tourist behaviors are segmented using K-means clustering (k=7) applied to user-POI interaction matrices, enabling cluster-specific prompting strategies.

- **Anchor-based Sequence Construction:** We implement a configurable anchor selection mechanism (`DEFAULT_ANCHOR_RULE = "penultimate"`) to determine the reference point for next-POI prediction within user trajectories.

- **Distance-aware POI Ranking:** Geospatial prompts incorporate Haversine distance calculations to rank available POIs by proximity to the current location, reflecting realistic mobility constraints.

The prompting mechanism dynamically generates context-aware queries following this template structure:

Listing 1: Comprehensive Context Prompt Template

```
Sei un assistente turistico esperto di Verona con conoscenza
dettagliata della geografia della citt .

<cluster_id>: {cluster_id}
<history>: {history}
<current_poi>: {current_poi}

<pois_disponibili_con_distanze>:
{pois_with_distance_text}

Obiettivo: suggerisci i {top_k} POI pi   probabili che
l'utente visiter   dopo, considerando:
– La distanza dal POI attuale
– La logica dei percorsi turistici a Verona
– I pattern tipici di movimento in base al cluster {cluster_id}
```

4.3 Evaluation Metrics

We employ a comprehensive evaluation framework encompassing multiple recommendation quality metrics:

- **Top-1 Accuracy:** $\mathrm{Acc}_{@1} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{y_i = \hat{y}_i^{(1)}\}$

- **Top-k Hit Rate:** $\mathrm{HR}_{@k} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{y_i \in \{\hat{y}_i^{(1)}, \dots, \hat{y}_i^{(k)}\}\}$

- **Mean Reciprocal Rank:** $\mathrm{MRR} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\mathrm{rank}_i}$

- **Catalogue Coverage:** $\mathrm{Coverage} = \frac{|\bigcup_i \{\hat{y}_i^{(1)}, \dots, \hat{y}_i^{(k)}\}|}{|\mathcal{P}|}$

where y_i represents the ground-truth next POI, $\hat{y}_i^{(j)}$ denotes the j-th ranked prediction, and $\mathcal{P}$ is the complete POI catalogue.

4.4 Experimental Results

Our experimental evaluation reveals significant performance variations across the three prompting strategies, with notable improvements observed when incorporating geospatial and temporal contextual information.

4.4.1 Cross-Strategy Performance Comparison

The progressive enhancement of prompting strategies demonstrates clear performance improvements:

4.4.2 Temporal Performance Analysis

The longitudinal analysis across the dataset timespan (2014-2020) reveals temporal stability in model performance, with year-over-year consistency in prediction accuracy metrics.

Table 1: Performance Comparison Across Prompting Strategies

Strategy	Top-1 Accuracy	Top-5 Hit Rate	MRR
POI Name Only	–	–	–
Name + Geolocation	–	–	–
Name + Geolocation + Temporal	–	–	–

Figure 1: Performance metrics by year showing temporal stability across the evaluation period.

4.4.3 Geographical Context Impact

The incorporation of geospatial information demonstrates substantial improvements in recommendation quality. Distance-aware prompting enables the model to capture realistic mobility constraints, with tourists showing strong preference for geographically proximate POIs.

Figure 2: Impact of geographical context on recommendation accuracy, showing performance improvements with distance-aware prompting.

4.4.4 Cluster-Specific Performance

The user segmentation approach reveals heterogeneous performance patterns across different tourist behavioral clusters, suggesting that personalized prompting strategies may further enhance recommendation quality.

4.5 Error Analysis and Model Interpretability

To understand the limitations and failure modes of our approach, we conduct a comprehensive error analysis examining:

- **Worst-performing POI pairs:** Identification of systematic prediction errors between specific POI transitions

- **Confusion matrix analysis:** Detailed examination of prediction patterns for the most frequently visited POIs

- **Explainability through LIME:** Text-based feature importance analysis to understand which contextual elements drive specific predictions

The error analysis reveals that geographic proximity is the dominant factor in successful predictions, with failures often occurring when tourists deviate from expected spatial clustering patterns. The LIME-based explainability analysis demonstrates that cluster membership and historical visitation patterns are key contextual features influencing model decisions.

Figure 3: Performance breakdown by user cluster, demonstrating the effectiveness of behaviorally-informed prompting strategies.

4.6 Discussion

The experimental results demonstrate that contextual enrichment significantly enhances LLM performance in next-POI prediction tasks. The geospatial-enhanced strategy shows the most substantial improvement over the baseline, while temporal context provides additional but more marginal gains. This suggests that spatial proximity constraints are the primary drivers of tourist mobility behavior, with temporal patterns playing a secondary role.

The cluster-specific analysis reveals the importance of user behavioral segmentation in achieving optimal recommendation performance. Different tourist types exhibit distinct mobility patterns, suggesting that adaptive prompting strategies could further improve prediction accuracy.

These findings have important implications for tourism recommendation systems, demonstrating that LLMs can effectively capture complex spatio-temporal mobility patterns when provided with appropriately structured contextual information.

5 Experimental Setup

5.1 Evaluation Methodology

We evaluate the LLM-Mob framework using a comprehensive experimental design that captures different aspects of mobility prediction performance:

5.1.1 Test Set Construction

For each tourist with at least three POI visits, we construct prediction instances by:

1. Taking the full visit sequence $S = \{p_1, p_2, ..., p_n\}$

2. Using $\{p_1, p_2, ..., p_{n-1}\}$ as input context

3. Predicting p_n as the target

5.1.2 Anchor Point Selection

We implement flexible anchor point selection strategies to investigate the impact of different reference points:

- **Penultimate**: Uses the second-to-last visited POI as the current location

- **First**: Uses the first visited POI as the current location

- **Middle**: Uses the middle POI in the sequence as the current location

- **Custom Index**: Allows specification of arbitrary anchor points

5.1.3 Performance Metrics

We evaluate prediction performance using Hit@K accuracy, where a prediction is considered correct if the ground truth POI appears in the top-K predicted locations. Our primary focus is on Hit@5 accuracy to capture the practical utility of the system.

5.2 Experimental Configurations

Our experiments are designed to investigate the following research questions:

1. How does the inclusion of geographical information affect prediction accuracy?

2. What is the impact of temporal context on mobility prediction performance?

3. How do different tourist behavioral clusters influence prediction outcomes?

4. What role does geographical proximity play in tourist movement patterns?

6 Results and Analysis

6.1 Performance Comparison Across Context Types

Table 2: Prediction accuracy comparison across different contextual information types

Context Type	Hit@1	Hit@3	Hit@5
POI Names Only	-	-	-
POI Names + Geography	-	-	-
POI Names + Geography + Temporal	-	-	-

6.2 Impact of Geographical Information

The integration of geographical information represents a significant advancement in our mobility prediction framework. Our analysis reveals several key insights:

6.2.1 Distance-Based Predictions

The inclusion of geographical coordinates and distance calculations allows the LLM to make more informed predictions based on spatial proximity. Tourists typically exhibit a preference for visiting nearby attractions, and our geographical context enables the model to capture these patterns effectively.

6.2.2 Spatial Clustering Analysis

We observe distinct spatial clustering patterns in tourist behavior, with certain POI combinations showing higher co-occurrence rates. The geographical context helps the LLM identify these patterns and make predictions that align with typical tourist routing behaviors.

Figure 4: Geographical distribution of POI predictions and their accuracy rates

6.3 Temporal Pattern Analysis

The incorporation of temporal information provides additional context that enhances prediction accuracy:

6.3.1 Time-of-Day Effects

Our analysis reveals that visiting patterns vary significantly based on the time of day, with certain POIs being more popular during specific time periods. The temporal context enables the LLM to capture these nuanced patterns.

6.3.2 Seasonal Variations

The multi-year dataset allows us to analyze seasonal effects on tourist behavior, revealing preferences for certain attractions during different times of the year.

Figure 5: Temporal patterns in tourist mobility and their impact on prediction accuracy

6.4 Cluster-Based Behavior Analysis

The K-means clustering approach reveals distinct tourist profiles that influence mobility patterns:

6.4.1 Cultural Tourists

One cluster exhibits strong preferences for cultural and historical sites, with predictable patterns of movement between related attractions.

6.4.2 Casual Visitors

Another cluster shows more diverse visiting patterns, suggesting casual tourists with varied interests.

6.4.3 Systematic Explorers

A third cluster demonstrates systematic exploration patterns, visiting attractions in geographical proximity.

Figure 6: Tourist cluster characteristics and their impact on mobility prediction

7 Discussion

7.1 Advantages of LLM-Based Approach

The LLM-Mob framework offers several advantages over traditional mobility prediction approaches:

7.1.1 Reduced Feature Engineering

Unlike traditional machine learning approaches that require extensive feature engineering, our LLM-based method leverages natural language descriptions to encode complex spatial-temporal relationships.

7.1.2 Contextual Understanding

LLMs demonstrate inherent capabilities for understanding contextual relationships between locations, enabling more nuanced predictions that consider factors beyond simple spatial proximity.

7.1.3 Interpretability

The natural language output of LLMs provides interpretable explanations for predictions, offering insights into the reasoning behind mobility choices.

7.2 Limitations and Challenges

Despite its promising results, the LLM-Mob framework faces several limitations:

7.2.1 Computational Requirements

Running LLMs locally requires significant computational resources, particularly for GPU-accelerated inference.

7.2.2 Prompt Sensitivity

The performance of LLM-based predictions can be sensitive to prompt design and formatting, requiring careful engineering of input contexts.

7.2.3 Scalability Concerns

Processing large datasets with LLMs can be time-consuming, particularly when compared to traditional machine learning approaches.

8 Future Work

Several directions for future research emerge from this work:

8.1 Multi-Modal Integration

Future work could explore the integration of additional modalities, such as images of POIs or textual descriptions of attractions, to provide richer context for mobility prediction.

8.2 Real-Time Prediction

Developing real-time mobility prediction systems that can adapt to changing conditions and provide dynamic recommendations to tourists.

8.3 Cross-City Generalization

Investigating the generalizability of LLM-based mobility prediction across different cities and cultural contexts.

8.4 Personalization

Exploring personalization techniques that can adapt predictions to individual tourist preferences and behaviors.

9 Conclusion

This paper presents the power of general (non-fine-tuning) Large Language Models (LLMs) for predicting human mobility in the city of Verona.
Using the VeronaCard dataset, we demonstrate the potential of LLMs to understand complex spatiotemporal patterns of tourist behavior without requiring in-depth feature engineering.

The analysis of different types of contextual information reveals the importance of the prompt, and especially the geographic and temporal context, in mobility prediction. The framework's use of open-source models (such as Llama) makes it accessible to researchers and practitioners, while maintaining competitive performance.

Our results contribute to the growing body of work on LLM applications in spatiotemporal analysis and provide the foundation for future research on intelligent tourism systems and urban mobility prediction.

10 Implementation Details

The complete implementation of thi paper is available as open-source software on Git-Hub repository at: `https://github.com/simo-hue/LLM-Mob-As-Mobility-Interpreter.git`
The repository includes:

- Implementation of the LLM-Mob framework without the OpenAI API dependencies

- Complete Python implementation with local LLM inference (For Example With Llama 3.1 integration)

- Data preprocessing pipelines for the VeronaCard dataset

- Experimental scripts (Notebooks) for evaluate the results

- instructions for setting up the environment and running the experiments

- Documentation and setup instructions (in addition to the readme.md file)

11   Acknowledgments

I would like to begin by expressing my deepest gratitude to my thesis advisor, Sara Migliorini, whose support, expert guidance, and constructive feedback have been crucial throughout this research/study. I am also grateful to the University of Verona for providing me with access to the valuable VeronaCard dataset, which served as the basis for the experimental validation presented in this study.

Special thanks go to CINECA for granting me access to the Leonardo supercomputer through the ISCRA (Italian SuperComputing Resource Allocation) initiative. I thank CINECA for the ISCRA award, for the availability of high-performance computing resources, and for the associated support. The computational experiments conducted in this research would not have been (so easily) possible without the exceptional performance of the Leonardo Booster partition. The high-performance computing environment offered by CINECA was essential to enabling the comprehensive experimental validation and achieving the results presented in this thesis.

I am particularly grateful for the opportunity to use such a cutting-edge computational infrastructure, which not only accelerated the research process but also allowed me to explore more sophisticated experimental designs and conduct more in-depth analyses than would have been possible with conventional computing resources. The technical support and documentation provided by the CINECA team were invaluable in navigating the complexities of high-performance computing environments.

Finally, I express my gratitude to all the colleagues, friends, and family who have supported me throughout this academic journey, offering encouragement and understanding during the most challenging phases of this research. Their moral support has been an essential component of my ability to successfully complete this work.