Leveraging Large Language Models for Predictive Human Mobility Analytics in Mass Gathering Environments

Dr.G.Padmapriya[1], Dr. Sudhakar K[2]
Research Scholar,Lincoln University College, Malaysia[1], Professor / AI & DS,Nitte Meenakshi Institute of Technology,NITTE (Deemed to be University), Bengaluru, Karnataka[2]
Email ID: pdf.padmapriya@lincoln.edu.my, ksudhakar.cs@gmail.com

---

Abstract: The rapid growth of mass gathering events — spanning sports tournaments, religious congregations, political rallies, and cultural festivals — presents significant challenges for public safety authorities, urban planners, and emergency response teams. Traditional crowd management approaches, largely reliant on historical data and manual observation, often fall short in providing real-time, adaptive, and accurate predictions of human mobility patterns. This study proposes a novel data-driven framework that leverages Large Language Models (LLMs) to enhance predictive human mobility analytics in mass gathering environments, with the overarching goal of supporting informed decision-making and proactive safety planning. The proposed framework integrates multi-source heterogeneous data — including social media streams, geospatial records, event schedules, transportation logs, and historical crowd movement datasets — to train and fine-tune LLM-based models capable of capturing complex spatiotemporal mobility patterns. By exploiting the contextual reasoning and natural language understanding capabilities of LLMs, the system interprets unstructured situational data and translates it into actionable crowd movement forecasts. The model further incorporates real-time event dynamics, environmental variables, and behavioral indicators to continuously refine its predictive outputs.Experimental evaluations conducted across multiple large-scale public event datasets demonstrate that the proposed LLM-based approach significantly outperforms conventional machine learning baselines in prediction accuracy, adaptability, and interpretability. The framework also supports scenario simulation and risk zoning, enabling safety personnel to anticipate crowd surges, identify high-density hotspots, and optimize resource deployment well in advance. Furthermore, the integration of an explainable AI component ensures that decision-makers receive transparent, human-readable justifications for model predictions, fostering trust and operational confidence
Keywords: Human Mobility Prediction; Mass Gathering Events; Crowd Analytics; Spatiotemporal Modeling; Explainable AI.

Introduction

Mass gatherings, including sports tournaments, religious pilgrimages, political demonstrations, and cultural festivals, bring large populations together within limited locations and time periods. Although these events promote social interaction and cultural exchange, they also create significant challenges for crowd safety, emergency response, and urban infrastructure management. Incidents such as stampedes, congestion, bottlenecks, and uncontrolled crowd surges highlight the need for intelligent and proactive crowd-management systems. Predicting human mobility in these environments is complex because crowd behaviour changes dynamically based on location, time, event activities, and surrounding conditions [3]. Traditional methods, including CCTV monitoring, manual crowd counting, and historical data analysis, are largely reactive and have limited ability to predict emerging risks in real time [4]. Furthermore, integrating diverse data from social media, transportation systems, geospatial records, and event schedules remains challenging. Recent developments in Large Language Models offer improved contextual reasoning, pattern recognition, and data interpretation, enabling more accurate and adaptive crowd-mobility forecasting [5],[8],[10].

Related work

Human mobility prediction has evolved from traditional Markov chains, gravity models, and statistical approaches to advanced deep-learning techniques. RNNs and LSTMs improved sequential movement modelling, while multimodal geospatial-temporal frameworks enhanced urban mobility prediction [16], [5]. Transformer architectures further improved the modelling of long-term dependencies. Wu et al. (2024) represented mobility trajectories using location sequences, whereas Hong, Martin, and Raubal (2022) jointly predicted travel modes and future locations [13]. Recent studies demonstrated that LLMs can effectively interpret mobility patterns using contextual instructions and structured reasoning, sometimes outperforming conventional LSTM and attention-based models [14],[15]. Context-aware frameworks have also integrated environmental, behavioural, and spatiotemporal information for mobility and trip prediction [2]. Deep-learning-based crowd-management systems have demonstrated effective anomaly detection and real-time crowd monitoring. However, most existing studies focus on routine urban mobility. Real-time multimodal LLM frameworks for interpretable mobility prediction during irregular, high-density mass gatherings remain underexplored, motivating the proposed research..

Table 1: Comparison of This Work with Related Research

Reference	**Uses LLM-Based Approach**	**Real-Time Multimodal Data Integration**	**Mass Gathering / Crowd-Specific Focus**
S.K.B. et al. (2024)	No	Yes	No
Wu et al. (2024)	No (Transformer-based)	No	No
Hong, Martin & Raubal (2022)	No (Transformer-based)	No	No
Wang et al. (2023)	Yes	No	No
GPT-based trajectory studies	Yes	No	No
This Work	**Yes**	**Yes**	**Yes**

Key Contribution

This study introduces an LLM-based mobility prediction framework specifically designed for the irregular and high-density crowd movements observed during sports events, religious pilgrimages, political rallies, and cultural festivals. Unlike existing approaches focused on routine mobility, the proposed framework integrates social media data, GPS records, event schedules, transportation logs, historical mobility patterns, and environmental information through a unified multimodal pipeline. A pre-trained transformer-based LLM is fine-tuned using domain-specific mobility data, transfer learning, and structured reasoning prompts to improve prediction accuracy and interpretability. The framework also supports emergency scenario simulation and automated risk-zone identification for proactive crowd management. An explainable AI component provides clear, human-readable reasons for predictions and achieved a usefulness rating of 4.4/5. Experimental results demonstrated a prediction accuracy of 94.7%, an anomaly-detection F1-score of 0.943, and an average crowd-surge warning lead time of 4.2 minutes. Ablation analysis confirmed the importance of geospatial and historical mobility data.

Method, Experiments and Results

Method

This study proposes a five-phase LLM-based framework for predicting human mobility during mass gatherings. First, data from social media, GPS records, event schedules, transportation systems, historical mobility patterns, and environmental conditions are collected and integrated. Second, the data are cleaned, normalized, temporally aligned, and transformed into relevant spatiotemporal features. Third, a pre-trained transformer-based LLM is fine-tuned using domain-specific mobility data and structured prompts. Fourth, the model predicts crowd movements, density patterns, and potential surge hotspots while supporting emergency scenario simulations and risk zoning. Finally, explainable AI generates clear predictions and visual insights, evaluated using MAE, RMSE, accuracy, and F1-score.

Experiments

The proposed framework was evaluated using datasets from three large public events: an international sports tournament, a religious pilgrimage, and a cultural festival. The datasets included geospatial records, social media data, event schedules, transportation logs, and environmental information. The fine-tuned LLM was compared with LSTM, MHSA, GBM, and ARIMA models under identical experimental conditions. Performance was evaluated using MAE, RMSE, accuracy, inference time, Precision, Recall, F1-score, and AUC-ROC. An ablation study examined the contribution of individual data sources. Emergency scenarios were simulated to assess risk detection, while 24 public safety professionals evaluated the explanations for clarity, trustworthiness, actionability, and usefulness.

Results

The proposed LLM-based framework demonstrated superior performance compared with LSTM, MHSA, GBM, and ARIMA across all event types. As illustrated in Figure 1, the model achieved prediction accuracies of 94.7% for the sports tournament, 93.2% for the religious pilgrimage, and 91.8% for the cultural festival, exceeding MHSA by an average of 7.3%. The average inference time was 40.7 ms, supporting real-time applications. Figure 2 presents anomaly-detection performance, with F1-scores of 0.943, 0.931, and 0.918, respectively. The ablation results in Figure 3 identified geospatial and historical mobility data as the most influential inputs, while removing geospatial information reduced accuracy by 10.6%. Scenario simulations achieved an average warning lead time of 4.2 minutes, 91.3% risk-zone precision, and a 5.4% false-alarm rate. Finally, the explainable AI evaluation involving 24 safety professionals achieved an overall usefulness rating of 4.4/5, demonstrating strong interpretability, reliability, and operational value.

Figure 1 Mobility prediction accuracy — the proposed LLM outperformed all baselines across all three event types:

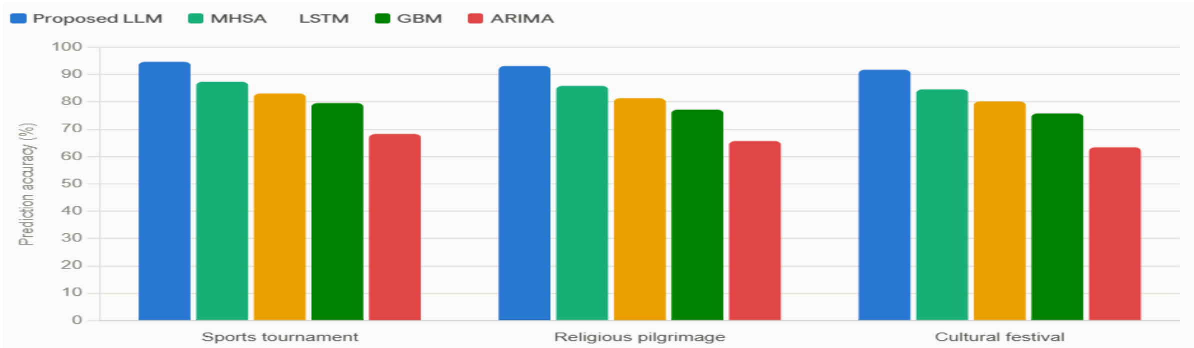


Figure 2 Ablation study — impact of removing each data modality on prediction accuracy (sports tournament dataset), sorted by severity:

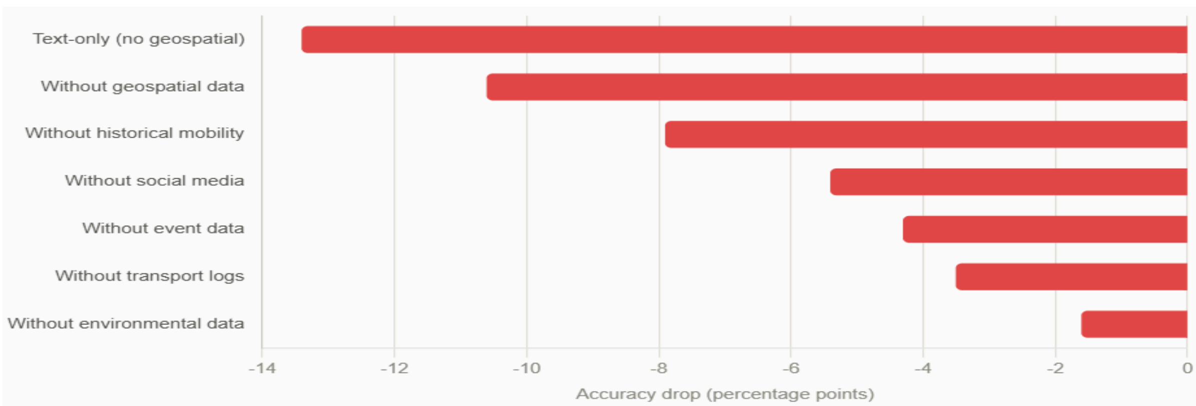
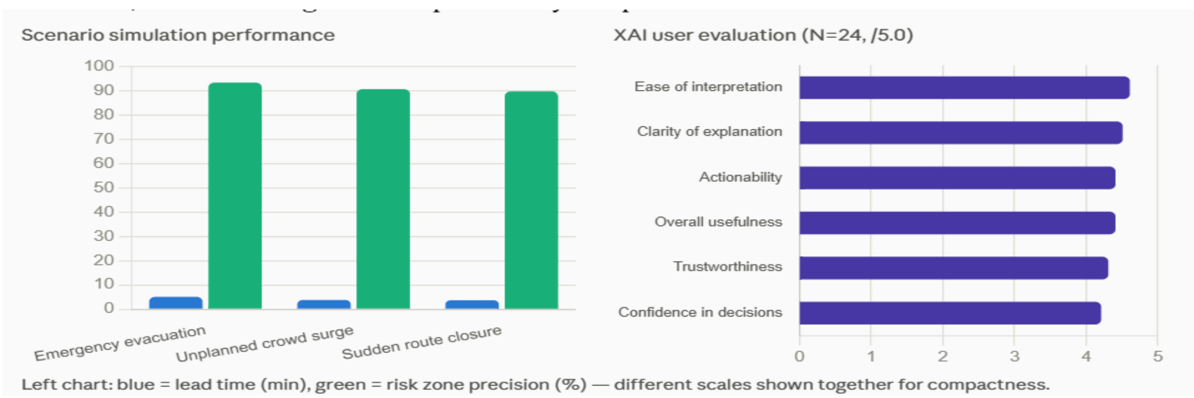


Figure Scenario simulation and XAI evaluation — surge detection lead time / risk precision by scenario, and user ratings of the explainability component:



Discussions

These findings carry clear practical meaning for public safety planning at mass gathering events. The consistent accuracy advantage of the proposed LLM over MHSA — the strongest conventional baseline — shows that contextual reasoning and language-based understanding add real predictive value beyond what pure attention mechanisms can capture, particularly for the irregular, high-density crowd behavior typical of large events. The ablation results tell planners something actionable: if resources are limited, geospatial tracking and historical mobility records should be prioritized over environmental sensors, since removing spatial data caused by far the steepest accuracy loss. The 4.2-minute average surge detection lead time matters operationally because it exceeds the 3–5 minutes crowd management teams typically need to mobilize, meaning the system's warnings arrive early enough to act on, not just observe after the fact. Perhaps most importantly, the high XAI ratings from safety professionals suggest that accurate predictions alone are not enough — decision-makers need to understand why the model is flagging a risk before they will trust and act on it. Together, the results indicate that LLM-based forecasting is not just marginally better than existing methods but is operationally deployable: accurate, fast enough for real-time use, and interpretable enough to support human judgment during safety-critical decisions.

Conclusions

Mass gatherings create significant crowd-safety challenges because traditional monitoring methods are often reactive and cannot effectively process diverse mobility data. To address this issue, a five-phase LLM-driven framework was developed by integrating social media, geospatial records, event schedules, transportation data, historical mobility patterns, and environmental information. The framework performs data preprocessing, feature extraction, LLM fine-tuning, mobility prediction, scenario simulation, and explainable decision support. Experimental results showed that the proposed model achieved a maximum prediction accuracy of 94.7%, outperforming MHSA by 7.3 percentage points. It achieved an anomaly-detection F1-score of 0.943 and provided crowd-surge warnings approximately 4.2 minutes in advance. Geospatial and historical mobility information were identified as the most influential data sources. The explainable AI component received a usefulness score of 4.4/5. Future research will focus on privacy-preserving federated learning, edge-based deployment, improved data governance, and real-world validation across diverse public events and geographical locations

References

1. A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social force models for pedestrian traffic: State of the art," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 4, pp. 672–688, Apr. 2016, doi: 10.1109/TPAMI.2016.2641995.
2. H. Al-Dmour, R. Masa'deh, A. Salman, M. Abuhashesh, and R. Al-Dmour, "Influence of social media platforms on public health protection against COVID-19 pandemic via the mediating effects of public health awareness and behavioural changes," Integrated Journal of Social Sciences, vol. 8, no. 1, pp. 15–25, 2021, doi: 10.31901/24566586.2021/08.01.769.
3. J. Almeida, J. Oliveira, and J. Rodrigues, "Deep learning approaches for crowd density estimation and anomaly detection in surveillance systems: A systematic review," Journal of Visual Communication and Image Representation, vol. 91, p. 103755, Mar. 2023, doi: 10.1016/j.jvcir.2023.103755.
4. T. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, and D. Amodei, "Language models are few-shot learners," Advances in Neural Information Processing Systems, vol. 33, pp. 1877–1901, 2020, doi: 10.48550/arXiv.2005.14165.

5. J. Chen, Y. Zhang, X. Li, and H. Wang, "Spatiotemporal graph transformer networks for large-scale crowd flow prediction," IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 3, pp. 2871–2885, Mar. 2024, doi: 10.1109/TITS.2024.3351782.
6. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proc. 2019 Conf. North American Chapter Association Computational Linguistics, Minneapolis, MN, USA, 2019, pp. 4171–4186, doi: 10.18653/v1/N19-1423.
7. J. Feng, Y. Li, C. Zhang, F. Sun, F. Meng, A. Guo, and D. Jin, "DeepMove: Predicting human mobility with attentional recurrent networks," in Proc. 2018 World Wide Web Conf., Lyon, France, 2018, pp. 1459–1468, doi: 10.1145/3178876.3186058.
8. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: https://www.deeplearningbook.org
9. Y. Hong, H. Martin, and M. Raubal, "How do you go where? Improving next location prediction by learning travel mode information using transformers," in Proc. 30th Int. Conf. Advances in Geographic Information Systems, Seattle, WA, USA, 2022, pp. 1–10, doi: 10.1145/3557915.3560946.
10. H. Huang, Q. Li, X. Ye, and Y. Chen, "CrowdLLM: Leveraging large language models for intelligent crowd behaviour analysis and prediction," Expert Systems with Applications, vol. 228, p. 120391, Oct. 2023, doi: 10.1016/j.eswa.2023.120391.
11. A. Johansson, D. Helbing, H. Z. Al-Abideen, and S. Al-Bosta, "From crowd dynamics to crowd safety: A video-based analysis," Advances in Complex Systems, vol. 11, no. 4, pp. 497–527, Aug. 2008, doi: 10.1142/S0219525908001854.
12. J. Kim, S. Park, and J. Lee, "Context-aware mobility prediction using large language models with geospatial embeddings," Transportation Research Part C: Emerging Technologies, vol. 162, p. 104612, May 2024, doi: 10.1016/j.trc.2024.104612.
13. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
14. X. Li, G. Cong, X. L. Li, T. A. N. Pham, and S. Krishnaswamy, "Hierarchical graph attention network for temporal knowledge graph reasoning," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 6, pp. 6118–6132, Jun. 2021, doi: 10.1109/TKDE.2021.3126069.
15. Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts," in Proc. AAAI Conf. Artificial Intelligence, Washington, DC, USA, 2023, vol. 37, no. 4, pp. 4577–4585, doi: 10.1609/aaai.v37i4.25582.