Journal of Transportation Engineering and its Applications
Volume 11, Issue 1, January-April 2026
ISSN: 3107-4499 (Online)

Large Language Models (LLMs) for Roadway Safety and Mobility Enhancement

Suman Diwaker[1], Sandeep Patil[2], Raju Joshi[3], Arvind Kotekar[4]

Associate Professor, Assistant Professor

Department of Transportation Engineering

CEPT University – Ahmedabad, Gujarat

Email: *Sumandiwaker293@gmail.com[1], patilsandeep09@rediffmail.com[2], rajujoshitui@yahoo.com[3]*

DOI: *https//doi.org/10.5281/zenodo.19216502*

ABSTRACT

Roadway safety and mobility are major concerns in rapidly urbanizing regions where traffic congestion, accidents, and inefficient transport systems are becoming serious problems. Recent developments in Artificial Intelligence have introduced Large Language Models (LLMs) as powerful tools capable of understanding, analyzing and generating human-like text from massive datasets. While LLMs are widely used in natural language processing applications, their potential role in transportation engineering, especially in roadway safety and mobility enhancement, is still emerging. This paper presents a comprehensive review of how LLMs can be applied in traffic accident analysis, traffic management, traveler information systems, infrastructure maintenance, and policy decision support. The paper also discusses integration of LLMs with IoT sensors, traffic cameras, GIS data, and intelligent transportation systems. Challenges such as data privacy, model bias, reliability and computational cost are also addressed. The study shows that LLMs can significantly support decision making, automate data interpretation, and improve safety strategies when combined with traditional traffic engineering tools.

KEYWORDS: *Large Language Models, Roadway Safety, Mobility Enhancement, Traffic Management, Intelligent Transportation Systems, AI in Transportation, Accident Analysis, Smart Mobility*

INTRODUCTION

Transportation systems are backbone of economic and social development. However, increase in vehicle population and urban expansion has caused serious roadway safety issues and mobility problems. Road accidents, congestion, delays and pollution are becoming everyday issues in cities as well as highways. Traditional traffic engineering methods rely heavily on numerical data and manual analysis which is often time consuming and limited in extracting hidden patterns from large unstructured data such as accident reports, driver complaints, CCTV footage descriptions and policy documents.

Large Language Models (LLMs) such as GPT, BERT, and PaLM have shown impressive ability in understanding context, extracting knowledge and generating meaningful interpretations from textual data. These models are trained on massive corpus and can process complex information in seconds. In roadway domain, huge amount of textual and semi-structured data exists which is usually underutilized.

This paper explores how LLMs can be utilized for improving roadway safety and mobility by transforming how traffic data, reports, and communication systems are analyzed and used.

OVERVIEW OF LARGE LANGUAGE MODELS (LLMS)

Large Language Models (LLMs) are advanced deep learning models built mainly on the **transformer architecture**, which was introduced to handle sequence-to-sequence learning without relying on recurrent structures. The key strength of transformers lies in the **self-attention mechanism**, through which the model can understand relationships between words in a sentence regardless of their distance from each other. Because of this, LLMs can capture context, meaning, and intent from large volumes of text data with very high accuracy.

LLMs are trained on extremely large text corpora collected from books, articles, websites, reports, and various digital sources. During training, the model learns language patterns, grammar, facts, reasoning styles, and contextual associations. As a result, LLMs are not limited to simple keyword matching; instead, they understand the **semantic meaning** of the content.

In transportation and roadway engineering, a large portion of available information is present in **unstructured or semi-structured textual form**, such as accident reports, citizen complaints, inspection notes, maintenance records, traffic police logs, CCTV event descriptions, and government policy documents. Traditional AI or machine learning techniques require this information to be converted into structured numerical format before analysis, which is time-consuming and often results in loss of valuable context. LLMs overcome this limitation by directly working with raw text data.

Transformer Architecture in LLMs

The transformer model consists of encoder and decoder layers with multiple attention heads. These attention heads allow the model to focus on different parts of a sentence simultaneously. For example, in an accident report sentence like:

"The collision occurred due to poor visibility caused by heavy rain near the curved section of the highway,"

the LLM can associate **poor visibility**, **heavy rain**, and **curved section** as related factors contributing to the accident. This level of contextual linking is very useful in identifying hidden patterns in roadway incidents.

Training and Fine-Tuning of LLMs

LLMs are first **pre-trained** on general language data and then can be **fine-tuned** on domain-specific datasets. For roadway applications, models can be fine-tuned using:

- Historical accident reports
- Traffic control room logs
- Road safety audit reports
- Transportation policy documents
- Public feedback records

This fine-tuning helps the model to understand domain-specific terminology such as *shoulder width, skid resistance, sight distance, signal phasing,* etc.

Key Characteristics of LLMs

The following capabilities of LLMs make them highly relevant for roadway safety and mobility studies:

- **Context Understanding**

  LLMs do not treat words independently. They understand the overall meaning of sentences and paragraphs. This helps in interpreting complex accident narratives and inspection notes.

- **Text Summarization**
  Long reports and policy documents can be summarized into short actionable points. Traffic engineers can quickly review large volumes of documentation.
- **Pattern Recognition from Reports**
  By reading thousands of accident reports, LLMs can identify repeating causes such as overspeeding near schools, poor lighting conditions, or faulty signage.
- **Question Answering and Decision Support**
  LLMs can be integrated into decision support systems where engineers or officials can ask questions like:
  "What are common causes of night-time accidents on highways?" and receive data-driven answers.
- **Multilingual Communication**
  In countries like India, traffic complaints and reports may be written in different languages. LLMs can translate and interpret multilingual data efficiently.
- **Knowledge Extraction from Large Documents**
  LLMs can extract specific information such as safety recommendations, compliance gaps, or maintenance needs from lengthy documents.

DATA SOURCES IN ROADWAY SYSTEMS SUITABLE FOR LLMS

Roadway and traffic management systems continuously generate a very large amount of data in day-to-day operations. A significant portion of this data is **textual, descriptive, or semi-structured**, which is usually not fully utilized in traditional traffic analysis. These data sources contain rich contextual information about accidents, congestion, infrastructure condition, driver behavior, and public perception. Large Language Models (LLMs) are particularly suitable for extracting meaningful insights from such data because they are designed to understand natural language in reports, logs, and documents.

Unlike numerical traffic counts or sensor readings, these textual records explain **why** an incident occurred, **how** people reacted, and **what** environmental or infrastructural factors were involved. This qualitative information is extremely important for roadway safety and mobility planning.

Accident and Incident Reports

Police departments, traffic control rooms, and hospitals generate detailed accident reports. These reports describe:

- Cause of accident
- Time and location
- Road and weather conditions
- Driver and vehicle behavior
- Injury severity and response time

These reports are usually written in narrative format. LLMs can read thousands of such reports and extract common patterns such as frequent causes, risky road segments, or time-based accident trends.

Traffic Police Logs and FIR Records

Traffic police maintain daily logs of violations, incidents, diversions, and enforcement activities. FIR (First Information Report) records often contain detailed descriptions of roadway events. LLMs can process these logs to identify:

- Areas with frequent violations
- Repeated rule-breaking behavior
- Ineffective enforcement zones

Citizen Complaints and Feedback Portals

Many smart cities and municipalities provide online portals and mobile apps where citizens report:

- Potholes and pavement damage
- Faulty traffic signals
- Missing signage
- Encroachment issues
- Unsafe pedestrian crossings

These complaints are highly valuable for identifying safety hazards. LLMs can cluster similar complaints and prioritize areas needing urgent attention.

CCTV Control Room Event Descriptions

While CCTV produces video data, operators often maintain text logs describing events such as:

- Vehicle breakdowns
- Road blockages
- Signal jumping
- Traffic congestion build-up

LLMs can analyze these logs to understand recurring mobility issues and incident patterns.

Road Safety Audit and Inspection Reports

Road safety audits, inspection reports, and field visit notes contain expert observations in textual form. These documents include comments on:

- Sight distance problems
- Shoulder conditions
- Signage visibility
- Lighting conditions
- Pedestrian facilities

LLMs can extract compliance gaps and recurring infrastructure issues from these documents.

Table 1: Roadway Data Types Usable by LLMs

Data Source	Description	Application
Accident reports	Police and hospital records	Accident pattern analysis
Traffic complaints	Citizen feedback	Identify unsafe zones
CCTV log descriptions	Incident narration	Real-time incident understanding
Maintenance logs	Road repair records	Infrastructure planning
Policy documents	Government guidelines	Policy evaluation

Data Source	Description	Application
Social media posts	Public alerts	Incident detection
Driver logs	Fleet reports	Driver behavior study

APPLICATIONS OF LLMS IN ROADWAY SAFETY

Accident Report Analysis

Thousands of accident reports are generated daily. Manual reading is difficult. LLMs can read reports and extract:

- Cause of accident
- Weather conditions
- Road geometry issues
- Driver behavior patterns

This helps in identifying accident blackspots faster.

Identification of Black Spots

By analyzing textual complaints, police reports and traffic violation logs, LLMs can cluster information and predict potential black spots before severe accidents occur.

Driver Behavior Understanding

Fleet operators maintain driver logs and feedback. LLMs can analyze this data to identify rash driving patterns and recommend training.

Road Safety Audits

LLMs can assist in reviewing road safety audit reports and highlighting missing compliance points with standards.

APPLICATIONS OF LLMS IN MOBILITY ENHANCEMENT

Traffic Congestion Analysis

LLMs can process traffic messages, control room logs, and commuter complaints to understand recurring congestion causes.

Journal of Transportation Engineering and its Applications
Volume 11, Issue 1, January-April 2026
ISSN: 3107-4499 (Online)

Intelligent Traveler Information Systems

LLMs can be used in chatbots to provide real-time route guidance and safety alerts to commuters in local languages.

Public Transport Feedback Analysis

Passenger feedback about buses, metro and shared mobility can be analyzed to improve scheduling and service quality.

Incident Management

During accidents or road blockages, LLM-based systems can summarize situation reports and assist traffic controllers in decision making.

INTEGRATION WITH INTELLIGENT TRANSPORTATION SYSTEMS (ITS)

LLMs alone are not sufficient. When integrated with ITS components like IoT sensors, GPS, GIS, and cameras, their efficiency increases.

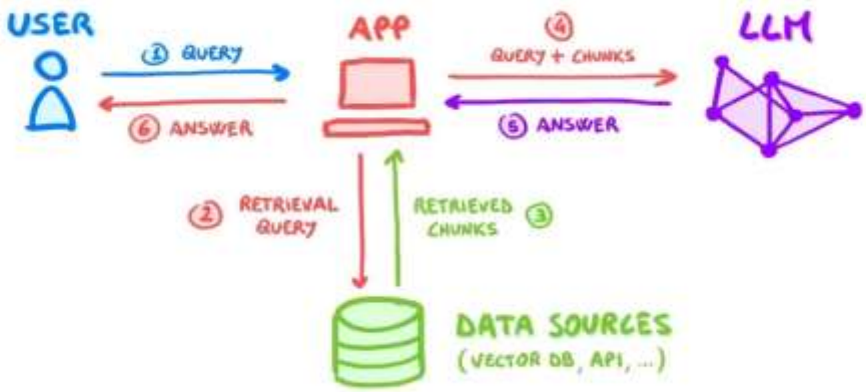

Figure 1: Integration of LLMs with ITS

LLMS FOR INFRASTRUCTURE MAINTENANCE

Maintenance logs contain history of potholes, pavement distress, signage damages etc. LLMs can analyze historical records and suggest preventive maintenance schedules.

ROLE IN POLICY MAKING AND PLANNING

Government agencies produce lengthy transport policies. LLMs can:

- Summarize policies

- Compare old and new policies
- Identify missing safety measures
- Assist planners in drafting new regulations

CASE STUDY EXAMPLE (HYPOTHETICAL)

In a mid-sized Indian city, accident reports of 5 years were processed using LLM. The model identified that majority accidents occurred near unmarked pedestrian crossings near schools during evening hours. Based on insights, signage and speed calming measures were installed, reducing accidents by 28% in one year.

BENEFITS OF USING LLMS IN ROADWAY SYSTEMS

- Faster data processing
- Reduced manual effort
- Better decision making
- Early risk detection
- Multilingual communication
- Knowledge extraction from legacy documents

CHALLENGES AND LIMITATIONS

- Data privacy concerns
- Model bias due to training data
- Requirement of high computing resources
- Need of domain specific fine-tuning
- Reliability and accountability issues

FUTURE SCOPE

LLMs can be combined with computer vision models for complete roadway intelligence. Voice-based traffic advisory systems, automated report generation, and smart governance systems will be possible.

DISCUSSION

Though LLMs are primarily text-based models, roadway systems have large dependency on textual data which was earlier neglected. Using LLMs, traffic engineers can convert this untapped data into actionable knowledge. It will not replace engineers but support them.

CONCLUSION

Large Language Models have significant potential to transform roadway safety and mobility enhancement. From accident analysis to policy drafting and traveler information systems, LLMs can provide valuable support.

Integration with existing ITS infrastructure can lead to smarter, safer and more efficient transportation networks. However, challenges related to ethics, privacy and accuracy must be carefully managed. With proper implementation, LLMs can become an important tool in future transportation engineering practices.

REFERENCES

1. Vaswani, A. et al. (2017). Attention is All You Need. *NeurIPS*.
2. Devlin, J. et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers. *NAACL*.
3. Brown, T. et al. (2020). Language Models are Few-Shot Learners. *NeurIPS*.
4. Li, Y., Zheng, Y. (2021). AI Applications in Intelligent Transportation Systems. *IEEE ITS*.
5. Chen, C. et al. (2022). Natural Language Processing for Traffic Incident Analysis. *Transportation Research C*.
6. Kumar, P., Singh, R. (2023). AI-based Road Safety Analytics using Text Mining. *Journal of Transportation Safety*.
7. Zhang, L. et al. (2022). Smart Mobility using AI and Big Data. *Elsevier Smart Cities*.
8. Mehta, S., Kulkarni, A. (2023). Accident Report Mining using NLP Techniques. *International Journal of Traffic Engineering*.
9. World Health Organization (2023). Global Status Report on Road Safety.
10. Sharma, V., Patel, D. (2024). Intelligent Decision Support Systems in Traffic Management. *Journal of Intelligent Transport Systems*.

Cite as:

Suman Diwaker, Sandeep Patil, Raju Joshi, Arvind Kotekar. (2026). Large Language Models (LLMs) for Roadway Safety and Mobility Enhancement. Journal of Transportation Engineering and its Applications, 11(1), 42-56.
https//doi.org/10.5281/zenodo.19216502