

GenAI-Powered Smart Traffic Assistant: A Multimodal AI Framework for Real-Time Traffic Prediction, Route Recommendation, and Explainable Urban Mobility

Naganaboyina Vinod Kumar, Ch. Hema Vardhan, Gattu Mohan Goud, M. Mohan Krishna Reddy, U. Devendra

Department of Computer Science and Engineering (AIML), Parul University, Vadodara, Gujarat, India

2303031460111@paruluniversity.ac.in, 2303031460129@paruluniversity.ac.in, 2303031460257@paruluniversity.ac.in,

2303031460050@paruluniversity.ac.in, 2303031460246@paruluniversity.ac.in

Project guidance: Prof. Ritu R. Agarwal | Prototype deployment: Streamlit web application

Abstract

Urban traffic management increasingly requires systems that can reason over heterogeneous, time-varying information rather than rely only on static maps or historical averages. This paper presents the GenAI-Powered Smart Traffic Assistant, a prototype intelligent transportation platform that combines real-time location, traffic and weather inputs, historical traffic data, machine-learning-based congestion prediction, computer-vision vehicle detection, route recommendation, alerts, explainable AI (XAI), and role-aware dashboards in a unified workflow. The system is derived from a semester project report and its associated prototype deployment. The design follows a layered client-server architecture with a web/mobile-oriented interface, application services, machine-learning components, and persistent traffic data. A structured development methodology is used to integrate data collection, preprocessing, prediction, recommendation, visualization, security, and maintenance. Rather than claiming benchmark accuracy that is not reported in the source project, the paper evaluates the prototype through functional coverage, architecture consistency, safety constraints, and explicit non-functional targets. The resulting framework provides a practical foundation for multimodal traffic assistance and a pathway toward live CCTV integration, IoT sensing, adaptive signal control, continuous model learning, and smart-city interoperability.

Keywords: artificial intelligence, intelligent transportation systems, traffic prediction, route recommendation, computer vision, YOLO, weather-aware analytics, explainable AI, GenAI, smart cities

I. INTRODUCTION

Traffic congestion is increasingly difficult to manage because urban mobility is shaped by interacting spatial, temporal, and environmental factors. Vehicle demand varies over time, incidents create non-recurring delays, weather changes driver behaviour and road capacity, and incomplete sensing can obscure the current state of a road network. The supplied project report identifies these limitations in conventional approaches and proposes an intelligent assistant that combines real-time GPS location, environmental APIs, historical information, machine-learning prediction, route analysis, and vehicle detection.

The central research problem addressed in this work is the translation of heterogeneous traffic evidence into actionable, understandable assistance for a commuter or traffic authority. The proposed assistant is designed to ingest traffic observations, weather conditions, location context, and video-derived vehicle information; transform them into analytical features; estimate congestion; recommend routes; and communicate alerts or explanations through a responsive interface. The system therefore sits at the intersection of traffic forecasting, computer vision, decision support, and conversational assistance.

Recent research reinforces the need for models that capture spatial dependencies and temporal dynamics rather than using a single independent time series. STGCN explicitly models traffic

as a graph with spatio-temporal dependencies [4], DCRNN treats traffic propagation as diffusion on a directed graph [3], Graph WaveNet learns adaptive spatial relations and long temporal dependencies [5], and GMAN uses multi-attention blocks for long-horizon prediction [6]. Survey work also shows a shift toward multimodal, data-rich, and interpretable intelligent transportation systems [9]–[12].

The contributions of this paper are fourfold: (1) a unified system formulation for combining traffic, weather, location and vision signals; (2) a practical architecture aligned with the project report and its diagrams; (3) an evaluation framework that distinguishes design targets from measured results; and (4) a research roadmap for IoT, smart signals, multimodal learning and continuously updated models. The result is intended as a reproducible project-to-paper translation rather than a claim of production-scale traffic control.

II. RELATED WORK AND RESEARCH GAP

A. Traffic Forecasting with Spatio-Temporal Models

Modern traffic forecasting treats the road system as a structured network. DCRNN models traffic diffusion using bidirectional graph random walks and a recurrent encoder-decoder, reporting improvements over classical baselines on public traffic datasets [3]. STGCN replaces recurrent operations with graph and temporal convolutions, emphasizing efficient learning of spatial and temporal correlations [4]. Graph WaveNet extends this direction by learning adaptive dependency structures and using dilated temporal convolutions to capture long-range patterns [5]. GMAN further introduces graph multi-attention for direct modelling of relationships across spatial locations and future time steps [6]. These studies motivate the project choice of a modular prediction layer that can evolve from simpler supervised ML toward graph-based and attention-based architectures.

B. Computer Vision for Traffic Sensing

Vision provides an important complement to aggregate traffic APIs and sensor feeds. YOLO reframed object detection as a single regression problem and demonstrated real-time inference characteristics suitable for live video processing [1]. SORT later showed that lightweight online tracking can be layered on top of detections for real-time multi-object association [2]. For the proposed system, YOLO-style detection is used as the basis for vehicle counting and traffic-video analysis, with future work targeting live CCTV feeds, stronger tracking, and edge deployment.

C. External Conditions and Explainability

Traffic models increasingly incorporate exogenous variables such as weather, calendars, incidents and events. Recent forecasting literature emphasizes multimodal preprocessing and the value of combining heterogeneous inputs rather than treating traffic as an isolated sequence [10]–[12]. Temporal Fusion Transformer demonstrates how attention and gating can support multi-horizon forecasting while exposing interpretable variable and temporal

importance [8]. Such ideas align with the project requirement for XAI-oriented explanations that identify influential factors rather than exposing only a congestion score.

D. Research Gap

The literature contains highly specialized forecasting, detection, signal-control, and routing solutions, whereas the supplied project focuses on an integrated assistant experience. The gap is therefore architectural and operational: a single user-facing pipeline must connect data collection, prediction, recommendation, explanation, alerts, security, and persistent history. The project report also emphasizes that traffic predictions are advisory and should not be treated as authoritative commands, which is an important safety constraint when AI outputs influence real-world travel decisions.

Feature	Conventional / isolated	Proposed assistant
Live traffic	Single-source	Traffic + weather + location
Prediction	History / fixed rules	ML; graph/attention ready
Routing	Map-centric	Traffic-aware routing
Vehicle CV	Separate tool	Integrated module
Explainability	Often absent	Feature-based XAI
Alerts	Static	Congestion / incident / route
History	Limited	Traffic + prediction history
Stakeholders	Driver focused	Commuter + admin + authority

TABLE 1: COMPARATIVE FEATURE MATRIX OF EXISTING APPROACHES AND THE PROPOSED ASSISTANT

III. SYSTEM OBJECTIVES AND REQUIREMENTS

The project report defines a mobile-style intelligent application intended to improve urban travel efficiency by predicting traffic conditions, incorporating weather, analyzing alternative routes, detecting vehicles from video, retaining prediction history, and presenting information through a simple interface. The system supports three broad stakeholder groups: commuters, administrators, and traffic authorities. These roles influence the data each user can access and the actions available to them.

A. Functional Requirements

The functional scope includes user authentication and authorization; traffic data collection and validation; ML-based traffic prediction; congestion analysis; route recommendation; real-time alerts; emergency traffic assistance; traffic reporting; and dashboard visualization. The system also includes an XAI layer intended to explain the main factors behind model outputs. These capabilities are connected rather than treated as independent screens.

B. Non-Functional Requirements

The SRS specifies a normal-operation response-time target in the 3-5 second range for user actions and prediction/route modules. It also specifies scalability, high availability, modular maintainability, secure communication, role-based access, encrypted sensitive data, password protection, and data backup/recovery. These are design targets from the project requirements rather than measured benchmark results.

C. Safety and Governance

Because a traffic assistant can influence travel behaviour, system outputs are explicitly advisory. Traffic data should be validated before inference, emergency alerts should activate only in critical conditions, and administrators should control sensitive settings. The proposed architecture therefore treats trust, explanation, authentication, and data security as first-class system requirements rather than optional additions.

IV. SYSTEM ARCHITECTURE

The system follows a layered client-server architecture. The presentation layer provides a responsive dashboard and mobile-style interaction surface. The application layer manages authentication, routing of API calls, traffic-data processing, alerts and recommendation logic. The machine-learning layer performs congestion prediction and vehicle-oriented analysis. The data layer maintains historical traffic records, user information, system logs and analytical outputs. Cloud infrastructure is intended to provide deployment scalability and recovery mechanisms.

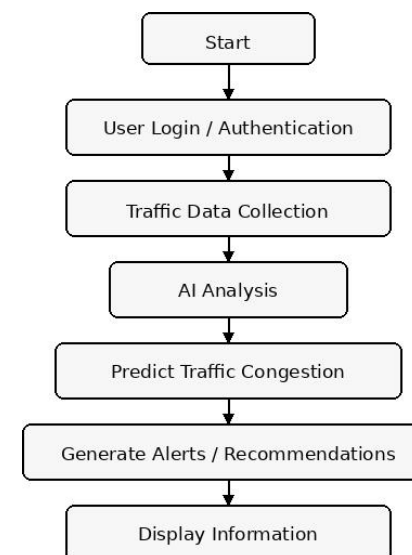


Fig. 1. End-to-end traffic-assistance pipeline derived from the project workflow and SRS.

The architecture deliberately separates acquisition from inference so that additional data sources can be introduced without redesigning the user interface. Traffic APIs, weather APIs, GPS context, and future IoT/CCTV feeds can be treated as sources that produce a normalized feature set. A prediction service can then consume this feature representation regardless of whether the underlying model is regression, tree-based ML, recurrent, graph-based, or attention-based.

V. DATA AND FEATURE DESIGN

The project report identifies four major data families: traffic observations, location context, environmental conditions, and historical prediction data. Traffic observations may include vehicle count, speed, congestion level, and road condition. Location context supplies the user position and route endpoints. Weather inputs include conditions such as rain, fog, temperature and other environmental indicators. Historical records provide temporal context for the learning and review cycle.

A practical normalized record can be represented as $x_t = [v_t, q_t, c_t, w_t, g_t, h_t]$, where v_t denotes speed-related evidence, q_t vehicle-volume evidence, c_t congestion context, w_t weather variables, g_t geographic context, and h_t historical/time features. The specific feature vector can be extended without changing the overall pipeline. This separation is

valuable because research-grade forecasting models often require consistent temporal alignment and careful handling of missing data.

The data layer also supports validation and secure storage. The report recommends PostgreSQL or MySQL for structured information, together with indexing, data-integrity controls, automated backup, and recovery. For a prototype-scale deployment, the same logical schema can be implemented with a simpler store and later migrated as data volume and concurrency increase.

Source	Key fields	Processing	Output
Traffic feeds	speed, volume, congestion	validate + normalize	traffic state
Weather APIs	rain, fog, temperature	align features	weather context
GPS / route	origin, destination, position	route context	candidate paths
Video / CCTV	frames, detections	YOLO + counting	vehicle density
History	past traffic + predictions	windows + trends	temporal context
User / system logs	preferences, alerts	security + audit	traceability

TABLE II: DATA SOURCES AND THEIR ROLE IN THE PROPOSED ANALYTICAL PIPELINE

VI. TRAFFIC PREDICTION AND RECOMMENDATION METHOD

A. Prediction Formulation

At time t , the assistant constructs an aligned feature vector from available traffic, weather, location and historical evidence. The prediction module estimates a future congestion state or traffic condition for one or more candidate road segments. The initial project description is intentionally model-agnostic; this is appropriate because the platform can begin with supervised ML and later adopt spatio-temporal graph networks as larger datasets become available. DCRNN, STGCN, Graph WaveNet and GMAN provide direct research directions for future model upgrades [3]–[6].

B. Route Scoring

Route recommendation combines predicted traffic state with travel distance and estimated delay. A conceptual route score can be written as $J(r) = \alpha T(r) + \beta C(r) + \gamma R(r)$, where T is predicted travel time, C is congestion burden and R represents route-risk or reliability penalties. The coefficients are tunable and are not reported as fitted values in the source project. The formulation clarifies how different objectives can be balanced without exposing model internals to the end user.

C. Weather-Aware Prediction

Weather is treated as an exogenous condition rather than a separate screen. Rain, fog and temperature can affect speed, visibility, road friction and driver behaviour. The literature increasingly supports multimodal fusion, and recent work explicitly studies weather-aware spatio-temporal forecasting [10], [15]. The proposed system therefore aligns weather observations to the same temporal and geographic context as traffic features before inference.

D. Computer Vision Module

The project uses YOLO-style object detection to detect vehicles from traffic video footage [1]. The resulting detections can be aggregated by class, lane or time interval to create density indicators. In later iterations, lightweight trackers such as SORT can be combined with detection to estimate motion and avoid repeated counting [2]. The project report positions this

component as a foundation for future live-CCTV integration rather than as a completed citywide vision deployment.

GENAI-POWERED SMART TRAFFIC ASSISTANT

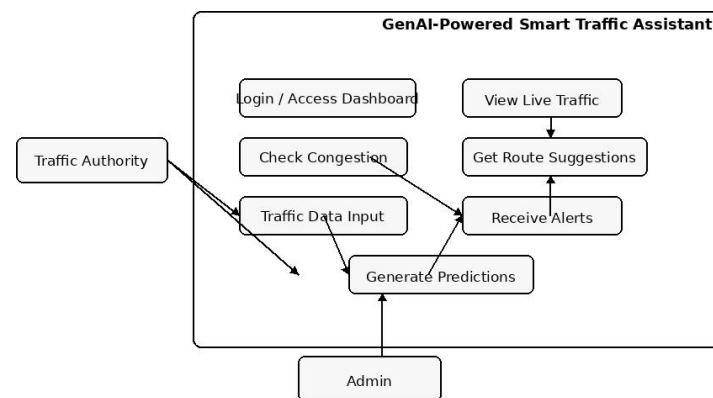


Fig. 2. Functional flow connecting users, prediction, alerts, route suggestions and administrative monitoring.

VII. EXPLAINABILITY, ALERTS, AND USER INTERACTION

The assistant is not designed merely to output a label such as “high congestion.” The report explicitly calls for explainable AI, so the interface should expose interpretable drivers such as rising vehicle volume, reduced speed, adverse weather, recent incident information, or historical peak-period behaviour. Temporal Fusion Transformer demonstrates one way of combining multi-horizon forecasting with feature-selection and attention-based interpretability [8]. In the proposed system, explanations are presented as concise user-facing rationales while the detailed model remains behind the service boundary.

The alert subsystem supports congestion notifications, route changes, travel-delay warnings, and emergency traffic events. Alerts are generated from prediction and rule checks, then delivered through the user interface. The emergency path should be conservative: uncertain detections should not trigger exaggerated instructions, and the UI should clearly distinguish advisory predictions from verified incidents.

The GenAI layer can be used as a natural-language presentation layer around structured outputs. Rather than allowing free-form generation to decide traffic actions, the assistant should retrieve verified traffic state, prediction summaries and route options from deterministic services and then generate concise explanations. This separation reduces the risk that language generation introduces unsupported traffic claims.

VIII. UML AND ROLE-BASED INTERACTION MODEL

The project report defines three main user roles: commuter, administrator and traffic authority. The UML view below translates those roles into a research-paper-oriented interaction model. Commuters log in, inspect traffic, check congestion, obtain route suggestions and receive alerts. Traffic authorities process and analyze traffic data, generate predictions, review reports and use system insights. Administrators maintain users, system operation and security controls.

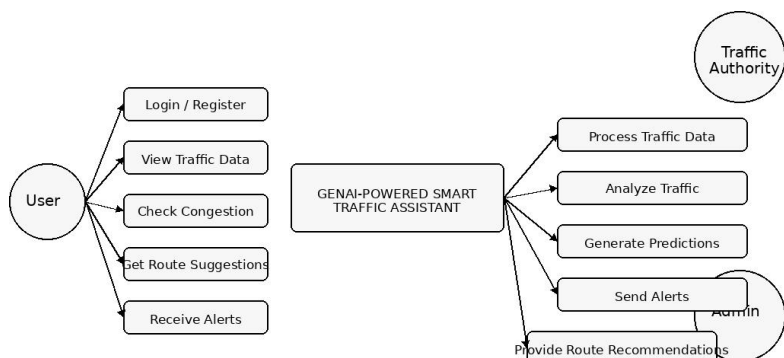


Fig. 3. UML-oriented interaction model redrawn from the project report for consistent paper presentation.

The role separation is important for security and deployment governance. Commuters require low-friction read access to traffic information, while authorities and administrators require broader analytical privileges. The project SRS therefore includes role-based access control, secure authentication, HTTPS communication, password hashing and protection of stored information.

IX. BASIC SYSTEM ARCHITECTURE AND DATA FLOW

The system flow begins with user authentication, proceeds through data collection and AI analysis, predicts congestion, generates alerts or recommendations, displays the result, and records relevant state back to persistent storage. This creates a closed operational loop: new observations inform predictions; predictions produce user actions; user interactions become part of the history that can support subsequent analysis.

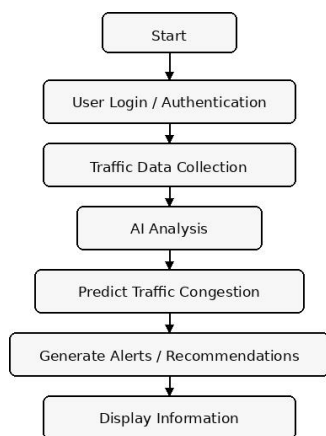


Fig. 4. Basic system architecture and processing sequence.

This sequence mirrors the project report's Waterfall-oriented implementation while keeping the deployed runtime architecture modular. In practice, data collection and prediction can execute asynchronously, allowing updates to arrive independently from the UI refresh cycle. The same architecture can later support message queues, edge preprocessing, model versioning and online monitoring without changing the core user workflow.

X. IMPLEMENTATION TECHNOLOGY STACK

Layer	Technologies in report	System role
Frontend	React; Tailwind CSS	dashboard + alerts
Backend	Node.js; Express.js	APIs + auth + logic
ML / analytics	Python; Scikit-learn; YOLO	prediction + detection
Database	MongoDB; PostgreSQL/MySQL option	history + users + traffic
Cloud	AWS / Azure; Docker-compatible	scale + availability
Realtime	WebSocket / API updates	live refresh + alerts

TABLE III: IMPLEMENTATION TECHNOLOGY STACK CONSOLIDATED FROM THE SOURCE PROJECT REPORT

XI. EVALUATION AND QUALITY ASSURANCE

The source project does not report a controlled benchmark study with a fixed public dataset, trained-model accuracy, or statistically replicated experiments. To avoid fabricating performance, this paper evaluates the design at the system level: functional completeness, architecture traceability, security requirements, responsiveness targets, and readiness for future quantitative validation.

A. Functional Test Matrix

Test area	Expected behaviour	Evidence	Status
Authentication	Authorized roles only	SRS security	Design
Traffic input	Validate before inference	SRS data rules	Design
Prediction	Return congestion insight	Prediction module	Prototype
Routing	Rank traffic-aware routes	Route requirements	Prototype
Weather	Weather informs analysis	Project scope	Prototype
Alerts	Deliver congestion / incident alerts	Alert requirements	Prototype
XAI	Show influential factors	XAI requirements	Design target
Backup	Restore critical data	DB requirements	Deployment

TABLE IV: SYSTEM-LEVEL EVALUATION MATRIX

The main measurable performance requirement supplied by the SRS is normal-operation response time of approximately 3-5 seconds for user interactions and prediction/route generation. This is treated here as a target service-level criterion, not as an observed benchmark. A publication-quality experimental evaluation should record latency distributions, prediction MAE/RMSE, precision/recall for vehicle detection, alert delivery time, route-delay reduction and user-level satisfaction under controlled scenarios.

B. Validation Strategy for a Next Research Iteration

A stronger empirical evaluation should use a public traffic benchmark such as METR-LA or PEMS-BAY for forecasting research, plus project-specific weather and route context. Metrics should include MAE, RMSE and MAPE for prediction, precision/recall/mAP for vehicle detection, route travel-time error, end-to-end latency and availability. Model ablations should

compare traffic-only features against traffic-plus-weather and traffic-plus-vision variants. Explainability should also be assessed through faithfulness and user comprehension rather than only visual inspection.

C. Research Reproducibility

The project architecture supports reproducibility because inputs, transformations, model versions and outputs can be persisted. Each prediction can be associated with a timestamp, route, input snapshot and model version. This makes later error analysis possible and allows researchers to determine whether failures arise from sensing, preprocessing, model drift, route data, or user-interface communication.

XII. LIMITATIONS AND DISCUSSION

The first limitation is data scale. The project report describes real-time and historical inputs but does not document a large standardized training corpus or a production-grade data-collection campaign. The second limitation is the absence of statistically reported model results, which limits claims about predictive superiority. The third limitation is dependence on external APIs and network connectivity. Real-time weather and map data may be incomplete, delayed or inconsistent with local road conditions.

A further limitation concerns computer vision under difficult conditions. Low light, occlusion, camera placement and dense traffic can reduce detection reliability. Finally, generative language interfaces require careful grounding: a language model should summarize verified model outputs, not invent traffic incidents or route states. These constraints are consistent with research findings that emphasize data quality, interpretability, spatial-temporal dependencies and deployment considerations in traffic prediction [9]-[12].

XIII. SMART-CITY INTEGRATION AND FUTURE WORK

The project report identifies a clear roadmap for extending the assistant beyond a prototype. First, IoT devices and smart sensors can enrich traffic state estimation with vehicle speed, density and environmental measurements. Second, live CCTV streams can move the vision module from offline clips toward continuous monitoring. Third, adaptive signal control can be incorporated through reinforcement learning after an appropriate safety and simulation layer is established. Research surveys show that DRL has become an active approach for traffic signal control, but also identify complexity and deployment challenges [13], [14].

Fourth, multimodal transportation can be introduced by incorporating buses, bicycles and pedestrian movement. Fifth, the route service can interoperate directly with navigation platforms. Sixth, continuous learning can update the model using newly arrived observations, while model governance prevents unvalidated updates from being deployed automatically. Recent work also points toward weather-aware multimodal spatio-temporal fusion and dynamic models that jointly encode external conditions [15].

GENAI-POWERED SMART TRAFFIC ASSISTANT

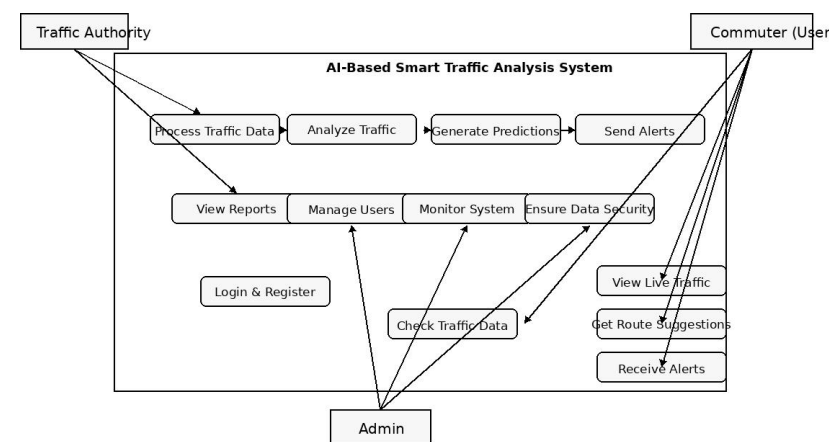


Fig. 5. Total use-case view of the proposed intelligent traffic assistant.

A mature deployment should also include a policy layer defining what the system may automate, what requires human verification, and which alerts are merely advisory. This distinction is essential when integrating with traffic authorities and smart-city control rooms. The assistant should support human decision makers rather than silently replace them in safety-critical functions.

XIV. CONCLUSION

This paper presented a GenAI-Powered Smart Traffic Assistant that integrates real-time traffic information, weather context, location, historical records, computer vision, congestion prediction, route recommendation, alerts, explainability and role-based monitoring in a unified architecture. The system translates the supplied project report into a research-oriented formulation and preserves the project's central design principle: heterogeneous traffic evidence should be converted into understandable, timely assistance rather than treated as isolated data feeds.

The work is best viewed as an integrated prototype framework with a clear path toward empirical research. Its main value lies in the architecture, data-flow organization and extensibility of the assistant. The next step is a controlled evaluation on public and locally collected datasets with reproducible model training, weather-aware ablations, route-level metrics, latency measurements and human-centered evaluation of explanations and alerts. With these additions, the project can evolve from a demonstration platform into a defensible intelligent transportation research system.

REFERENCES

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE CVPR, Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
- [2] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple Online and Realtime Tracking," in Proc. IEEE ICIP, 2016, pp. 3464-3468, doi: 10.1109/ICIP.2016.7533003.
- [3] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting," in Proc. ICLR, 2018.

- [4] B. Yu, H. Yin, and Z. Zhu, “Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting,” in Proc. IJCAI, 2018, pp. 3634-3640, doi: 10.24963/ijcai.2018/505.
- [5] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, “Graph WaveNet for Deep Spatial-Temporal Graph Modeling,” in Proc. IJCAI, 2019, pp. 1907-1913, doi: 10.24963/ijcai.2019/264.
- [6] C. Zheng, X. Fan, C. Wang, and J. Qi, “GMAN: A Graph Multi-Attention Network for Traffic Prediction,” in Proc. AAAI, vol. 34, no. 1, 2020, pp. 1234-1241, doi: 10.1609/aaai.v34i01.5477.
- [7] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, “Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting,” in Proc. AAAI, vol. 35, no. 12, 2021, pp. 11106-11115, doi: 10.1609/aaai.v35i12.17325.
- [8] B. Lim, S. O. Arik, N. Loeff, and T. Pfister, “Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting,” Int. J. Forecasting, vol. 37, no. 4, pp. 1748-1764, 2021, doi: 10.1016/j.ijforecast.2021.03.012.
- [9] N. Kumar and M. Raubal, “Applications of Deep Learning in Congestion Detection, Prediction and Alleviation: A Survey,” Transp. Res. Part C, vol. 133, p. 103432, 2021, doi: 10.1016/j.trc.2021.103432.
- [10] M. Shaygan, C. Meese, W. Li, X. G. Zhao, and M. Nejad, “Traffic Prediction Using Artificial Intelligence: Review of Recent Advances and Emerging Opportunities,” Transp. Res. Part C, vol. 145, p. 103921, 2022, doi: 10.1016/j.trc.2022.103921.
- [11] W. Jiang and J. Luo, “Graph Neural Network for Traffic Forecasting: A Survey,” Expert Syst. Appl., vol. 207, p. 117921, 2022, doi: 10.1016/j.eswa.2022.117921.
- [12] M. Afandizadeh et al., “Deep Learning Algorithms for Traffic Forecasting: A Comprehensive Review and Comparison with Classical Ones,” J. Adv. Transp., vol. 2024, Art. no. 9981657, 2024, doi: 10.1155/2024/9981657.
- [13] S. Bouktif, A. Cheniki, A. Ouni, and H. El-Sayed, “Deep Reinforcement Learning for Traffic Signal Control with Consistent State and Reward Design Approach,” Knowl.-Based Syst., vol. 267, p. 110440, 2023, doi: 10.1016/j.knosys.2023.110440.
- [14] F. Rasheed, K.-L. A. Yau, R. M. Noor, C. Wu, and Y.-C. Low, “Deep Reinforcement Learning for Traffic Signal Control: A Review,” IEEE Access, vol. 8, pp. 208016-208044, 2020, doi: 10.1109/ACCESS.2020.3034141.
- [15] G. Garbinčius and M. Makulavičius, “Urban Traffic Flow Simulation in Various Weather Conditions Using Long Short-Term Memory-Based Machine Learning,” Transport, 2026, doi: 10.3846/transport.2026.26773.

Source materials: supplied GenAI-Powered Smart Traffic Assistant project report; supplied AstraMedica reference PDF; project prototype URL supplied by the requester.