TrafficPulse: A Road-Sensor Assisted Traffic Tweet Misinformation Detection System

Junhao (Frank) Ran§, Yifan Wu§, Delaram Pirhayatifard, João Mattos, Arlei Silva
Rice University
Houston, TX, USA
{*jr151, evanwu, dp43, jrm28, arlei*}*@rice.edu*

***Abstract*—Traffic incident detection is a well-established task in transportation, traditionally addressed using a combination of traffic sensors and driver reports. More recently, social media has become a rich data source for timely incident detection. However, the highly dynamic nature of social media, the challenges in mapping textual content to precise real-world locations, and potentially misleading posts complicate the extraction of reliable traffic information. Motivated by these challenges and leveraging recent advances in large language models (LLMs), we propose a real-time tweet validation pipeline that extracts and verifies traffic incidents reported on X (former Twitter). Our approach employs advanced parsing techniques for localization extraction. It integrates publicly available data to confirm the existence of an incident, thereby enhancing the robustness of downstream traffic analysis methods that combine sensor data with verified textual features. To support further research in this domain, we also introduce two new datasets: the *Twitter Traffic Incidents* dataset, which comprises manually curated and human-verified incident reports, and the *PeMS Sensor + Incidents Reports* dataset, featuring snapshots from California's PeMS traffic sensor system. Experimental results demonstrate that our pipeline significantly improves the reliability of traffic incident validation in tweets, serving as a basis for future traffic anomaly detection research.**

***Index Terms*—Traffic Incident Detection, Social Media, Large Language Models, Traffic Sensors, Machine Learning**

1. Introduction

Traffic data analytics is a fundamental approach in transportation with direct societal implications for urban planning, public policy decision-making, and emergency response. By leveraging real-time data from traffic sensors, machine learning methods have shown great potential in addressing various tasks such as incident detection [1] , travel-time prediction [2], traffic flow prediction [3], and traffic speed prediction [4]. Moreover, in recent years, social media platforms, especially Twitter, have emerged as a valuable source for transportation-related data. The immediacy of social networks in reporting real-world events, combined with metadata such as geolocation and timestamps, offers features that can be integrated into traffic analytics models. Prior studies have demonstrated the positive impact of using textual features and temporal patterns for predicting traffic incidents, as well as next-morning congestion [5, 6, 7].

However, directly incorporating tweets into traffic prediction systems introduces significant challenges. The presence of fake or misleading reports in tweets can jeopardize the reliability of the entire prediction pipeline. This underscores the need for a robust mechanism for validating tweet-based data before its integration into traffic models.

Tweet validation is not a trivial task. The short contextual nature of tweets, combined with the potential for ambiguity and misinformation, makes veracity assessment a difficult problem. This paper aims to address this challenge while enhancing the availability of publicly available traffic and social media data for transportation research.

Our contributions are as follows:

- We demonstrate the use of large language model-based pipelines for geolocation extraction from tweet data.
- We propose a novel method for incident fact-checking by leveraging public data sources and machine learning approaches.
- We introduce two new datasets that link traffic sensor data and tweet-based incident reports, covering the state of California in 2017.

These contributions pave the way for more reliable integration of social media data into traffic prediction systems.

2. Related Work

The intersection of social media data mining and traffic analytics has garnered significant attention in recent years, with researchers leveraging the vast amount of user-generated content to enhance traffic monitoring and prediction. We provide an overview of the literature in this domain.

Social Media-based Traffic Event Detection Several studies have explored the potential of social media platforms, particularly Twitter, for real-time traffic event detection. In [8], it was proposed a system using Support Vector Machine (SVM) classification with a sigmoid kernel to detect traffic congestion in Jakarta, achieving an accuracy

§. Equal contribution

of 96.24%. Similarly, [9] developed a method for automatic detection of road traffic events from Arabic tweets using machine learning algorithms and the Apache Spark platform.

Gu et al. [6], propose to crawl, process, and filter tweets for extracting incident information on both highways and arterials. Their approach creates a dictionary of keywords implying traffic incidents and maps tweets into high-dimensional binary vectors for classification. Deep learning techniques, specifically Deep Belief Networks (DBN) and Long Short-Term Memory (LSTM) networks, were applied in [7] to detect traffic accidents from social media data, outperforming traditional methods like Support Vector Machines (SVMs) and supervised Latent Dirichlet Allocation.

Traffic Prediction and Analysis Beyond event detection, researchers have investigated the use of social media data for traffic prediction and analysis. In [10], the authors proposed a framework for next-day morning traffic prediction using Twitter data, leveraging people's work and rest patterns observed in evening and midnight tweets. Their approach outperformed existing methods, particularly for road segments with large day-to-day congestion variation.

A framework to identify and characterize accident- and congestion-prone areas discussed on social media, using natural language processing and deep learning techniques was proposed in [11]. Their study incorporated sentiment analysis to assess the priority level for mitigation measures, providing valuable insights for traffic authorities.

Fusion of Multiple Data Sources Recent work has explored integrating social media data with other data sources to enhance traffic monitoring capabilities. Daly et al. [12] proposed a method for fusing social media and linked data sources to understand real-time traffic conditions. Panakkal and Padgett [13] propose a framework leveraging physical, social, and visual sensors, along with physics-based models, to sense road conditions during flooding events. This approach demonstrates the potential of repurposing existing data sources to improve roadway situational awareness.

Classic Machine Learning Techniques Previous works [14] evaluate the performance of classic machine learning algorithms for real-time traffic classification using Twitter data, reporting SVM obtains the highest accuracy. However, deep learning models are still a promising way of capturing patterns in social media data for traffic-related tasks[7].

3. Background and Datasets

One of our contributions is the creation of two distinct datasets designed to support the verification pipeline. The first dataset consists of real traffic incidents reported by a government platform that collects data from traffic sensors. This dataset includes time-series data with real-time traffic measurements and recorded incidents. The second dataset is composed of tweets discussing traffic incidents, collected from Twitter during the same time periods as the governmental data. By pairing these datasets, we aim to provide a resource for tasks such as traffic incident prediction and extraction of traffic-related information from social media, advancing the state-of-the-art in both domains.

In this section we go over the construction details of both datasets. We also detail the annotation protocol of the tweets based on the obtained traffic incident reports and present an exploratory data analysis of these datasets.

3.1. PeMS and Traffic Measurements

Caltrans Performance Measurement System (PeMS)[1] is the data repository for the road system of the state of California([15]). PeMS collects 2 GB of traffic sensor data per day with 30-second snapshots in real-time. The sensor system has over 39,000 detectors on the freeways and aggregates the lane-by-lane value of traffic measurements such as flow, occupancy, and speed across all lanes. A group of sensors is called a *station*. Stations are responsible for collecting and centralizing the data of nearby sensors, defining a snapshot of a small segment of the transportation network. Stations are distributed along twelve districts.

PeMS also provides traffic-related incident records. There were approximately 519,532 incidents reported in 2017. The dataset includes comprehensive information about the incidents with a total of 20 attributes, including incident description, freeway number, and the specific geo-location. We match each incident reported with the closest station in a 3-mile radius, enabling a successful match for 470,041 (95%) of the incidents. The incidents are attributed to the snapshot of the station in the nearest hour of the timestamp of the incident, e.g. an incident reported at 10:43 am, will be assigned to the 11:00 am snapshot of that station.

Given the PeMS sensor system with $S := \{s_1, s_2, ..., s_N\}$ stations, our traffic dataset $\mathcal{D}_{\text{traffic}}$ consists of T timestamps organized in pairs $(x, y)_t^{s_i} \in \mathcal{D}_{\text{traffic}}^T$. Each pair (x, y), $x_t^{s_i} \in \mathbb{R}^M$ corresponds to a M-dimensional vector with the measurements captured by the sensors at station s_i at timestamp t, whereas $y_t^{s_i} \in \{0, 1\}$ is a label indicating the (in) existence of an incident in the timestamp.

Sensor measurements include the following:

- **Occupancy:** Proportion of time during an interval (e.g., 5 minutes) that a detection zone is occupied by vehicles, which serves as an indirect indicator of traffic density.
- **Flow:** Number of vehicles passing over a specific detector during a defined time interval. It is typically expressed in vehicles per time unit (e.g., 5-minute flow, hourly flow).
- **Speed:** Average speed (in mph) of vehicles passing during the defined time interval.

In our analysis, we consider a subset of the features available in the original dataset. The final dataset features include latitude, longitude, and separate flow, occupancy, and speed measurements per lane.

1. PeMS and the raw data can be accessed through https://pems.dot.ca.gov/.

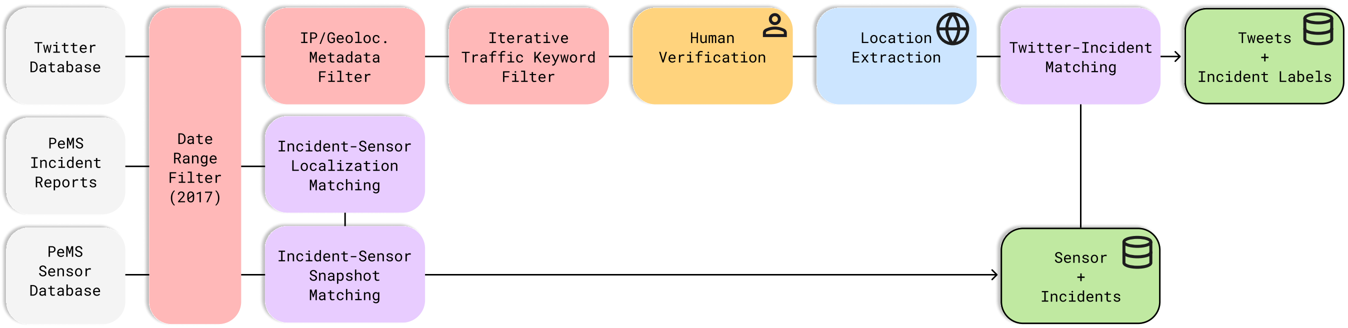

Figure 1. Overview of the dataset creation pipeline (Section 3). We use sensors to map filtered tweets to incidents, obtaining a label for each tweet.

TABLE 1. DESCRIPTION OF KEY TWITTER DATASET FIELDS

Field	Description
timestamp	UTC time when this Tweet was created.
text	The text of the tweet.
user	Properties of the creator of the tweet, e.g. the location and screen name, in JSON format
coordinates	Geographic location of this Tweet as reported by the user or client application.
place	When present, indicates that the tweet is associated (but not necessarily originating from) a Place \cite{twitterapidoc}. Provides rough geographical information compared to coordinates.

3.2. Traffic-Related Tweets

The Twitter Stream[2] is a public repository with dumps from Twitter from 2011 to 2023 in JSON format. Matching the period covered by the PeMS datasets, we extracted and focused our analysis on tweets from 2017. Each sample from Twitter Stream contains the text associated with each tweet, as well as the time and geographical information provided by the corresponding fields. In Table 1, we present the five key fields we considered for the localization and keyword filtering, the next steps in the dataset construction process.

To isolate tweets related to traffic incidents in California, we apply a set of filtering heuristics. For simplicity, we limit our analysis to English-language tweets, ensuring consistency in the keyword search. To filter by location, we use both the "location" field, when available, and the "coordinates" field to confirm that the tweet pertains to California. In cases where location data is missing or unclear, we check the user's location field (often indicating the user's IP address) or search the tweet text for references to California-related place names.

After applying location-based filters, the final step involves keyword-based filtering to identify traffic-related tweets. An initial set of keywords derived from domain knowledge is iteratively refined by inspecting the results of keyword searches. This iterative process expands the keyword set over three rounds, ensuring comprehensive coverage of relevant traffic incident topics. Appendix A provides the current set of traffic-related keywords.

Annotation and tweet validation We label each tweet extracted by verifying the existence of the incident reported through a standardized protocol:

1) A human reviewer assesses whether the filtered tweet explicitly reports a traffic-related incident in California. Only tweets with clear references to such incidents are retained.
2) The location described in the tweet is extracted.
3) The extracted location is then geocoded into coordinates (longitude and latitude).
4) Sensor data corresponding to the tweet's coordinates and timestamp is retrieved.
5) We consult the incident report system to verify whether the described event is listed in the official records. If the incident is reported, it is labeled as true; otherwise, it is labeled as false.
6) The final step involves annotating the tweet with a misinformation label. If the event is not found in the database, it is labeled as misinformation; otherwise, it is considered truthful.

This manual annotation process ensures high data accuracy by applying a verify-by-reference approach. This dataset enables the development and benchmarking of verification methods for incident reports based on social media posts, such as our proposed method.

3.3. Exploratory Data Analysis

We first analyze the impact of traffic incidents, as recorded in the Incident dataset, on key traffic parameters—average speed, occupancy, and flow—using data from nearby sensors in the PeMS dataset aggregated by hour. By correlating incident timestamps and locations with sensor data, we observe notable disruptions in traffic patterns following incidents.

Figure 4 highlight these effects. Specifically, average speed drops significantly after incidents, with a gradual recovery indicating the severity of the disruption. Occupancy rates spike post-incident, reflecting higher vehicle density, likely due to reduced speeds and potential lane closures.

2. https://archive.org/details/twitterstream

Figure 2. Wordcloud of Filtered Tweets

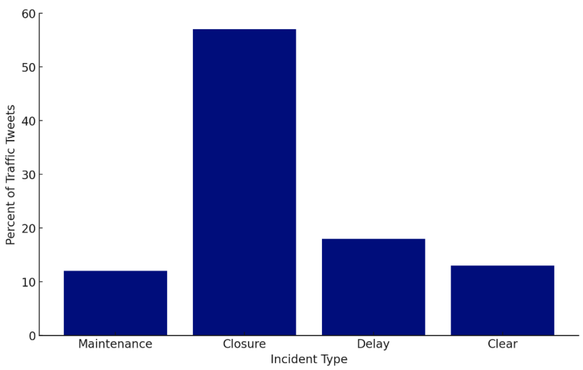

Figure 3. Incident Type Breakdown of Filtered Tweets

Additionally, the flow rate increases after the incident, suggesting a buildup of traffic passing through the sensors as vehicles recover from the incident's impact.

As a result of the tweet labeling pipeline, we investigate the nature of the incidents described in the filtered tweets. This step aims to align tweet data with traffic incident records and facilitate outlier detection within the incident data. The Incident dataset records all traffic-related incidents in all districts. The dataset includes all aspects of the incidents in detail, including description, freeway number, and the specific geo-location. The relevant fields are listed in the Table 3.

As illustrated in Figure 2, the most frequent terms in the filtered tweet dataset include *closed*, *maintenance*, and *traffic*, matching the keywords considered in the filtering step. Additional terms, such as *fires* and *tree*, suggest that the dataset encompasses a diverse range of events, provided they impact urban transportation. In addition, Figure 3 reveals that *closure* is by far the most frequently mentioned traffic-related incident in the tweets. This finding aligns with the prevalence of the term *closed* in the word cloud.

4. Traffic Misinformation Detection Pipeline

In this section, we present **TrafficPulse**, our real-time pipeline for detecting traffic-related misinformation on social media. TrafficPulse adopts a verify-by-reference strategy to assess the veracity of tweets by cross-referencing reported incidents with authoritative incident data. The

TABLE 2. DESCRIPTION OF KEY STATION 5-MINUTE DATASET FIELDS

Field	Description
Timestamp	The date and time of the beginning of the sensor data interval.
Station	A unique identifier for the station.
District	District number.
Freeway #	Freeway number.
Travel Direction	The direction of the traffic flow.
Total Flow	Sum of flows over the 5 minutes across all lanes.
Avg. Occupancy	Avg. occupancy across all lanes over 5 minutes expressed as a decimal number between 0 and 1.
Avg. Speed	Flow-weighted avg. speed over 5 minutes across all lanes.

TABLE 3. DESCRIPTION OF KEY STATION INCIDENT DATASET FIELDS

Field	Description
Timestamp	The date and time of the incident.
Description	A short description of the incident
District	District number.
Freeway #	Freeway number.
Location	The name of the road.
Latitude	The latitude coordinate of the incident.
Longitude	The longitude coordinate of the incident.

pipeline comprises three primary modules: (i) Tweet Reception, (ii) Traffic Information Extraction, and (iii) Incident Verification. Figure 5 provides an overview of the pipeline architecture.

Tweet Reception Module The first module is responsible for ingesting a high-volume stream of tweets in real-time. Given that tweets pertinent to traffic incidents are relatively scarce, we apply a targeted filtering algorithm based on a predefined set of traffic-related keywords (e.g., "crash," "collision," "incident" as detailed in Appendix B). In addition, tweets are filtered by geographic relevance to California using multiple cues, including user profile information, tweet content, and embedded geolocation metadata.

Traffic Information Extraction Module The second module refines the set of traffic-related tweets and extracts precise location details. This process is divided into two sub-components: Tweet Decision Tree and a geo-coding.

The Tweet Decision tree is implemented via ChatGPT-3.5-Turbo to evaluate and extract critical information from each tweet. At each decision node, the decision tree assesses the tweet against predetermined criteria structured as prompts, crafted through tweet patterns observed during the filtering procedure, described in Subsection 3.2. The prompts used in the decision process can be found in the Appendix B:

The geo-coding sub-component converts the extracted location descriptors into precise geographic coordinates. We adopt the ArcGIS Geocoding Service API. ArcGIS was chosen because of its robust support for geo-coding road intersections, a standard format found in our dataset.

Incident Verification Module The final module verifies the authenticity of reported incidents by matching the extracted geolocation and timestamp with records in an authoritative

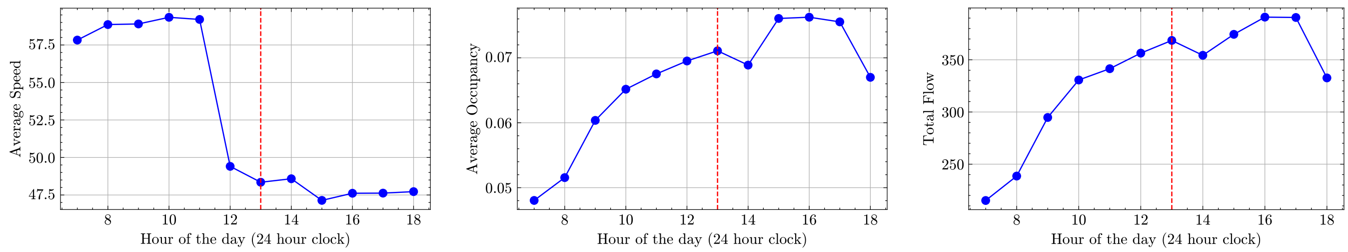

Figure 4. Left to Right - Impact of Incidents (marked in **red**) on Average Speed, Occupancy Rate, and Flow Rate.

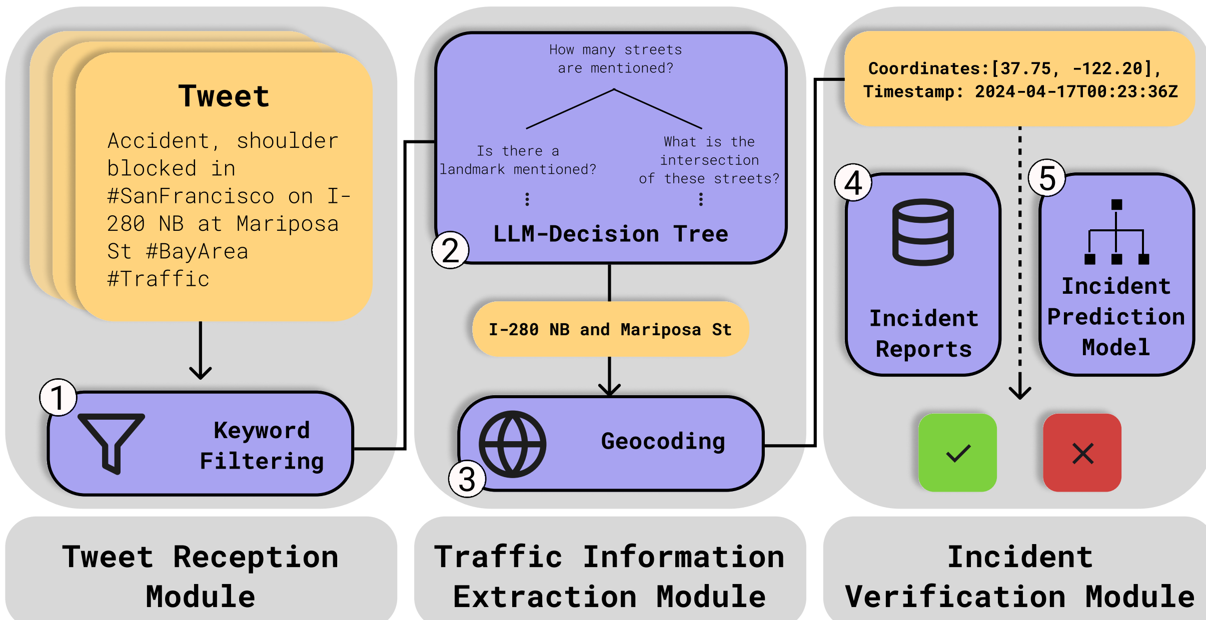

Figure 5. Overview of the Traffic Misinformation Detection Pipeline (Section 4). 1 - Keyword-based filtering ensures only tweets mentioning California and Traffic pass through the filter. 2 - Extracts relevant landmarks, streets, and freeway references for geocoding. 3 - Geocoding attempts to map a textual description to real-world coordinates through classical heuristics. 4 - Through the extracted coordinates and the timestamp of the tweet we can map the tweet to a snapshot in the official PeMS incident records, resulting in a ”truthful” tweet report if a match is found, and ”misleading” otherwise. 5 - Besides making use of the incident report databases, our pipeline output enables further analysis to be conducted by downstream machine learning methods by combining the extracted tweet incident report metadata with the sensor readings of the snapshots that match the timestamp of the tweet.

incident database. This module is designed to operate within a modular framework, allowing for the integration of machine learning–based incident prediction models in regions lacking rapid incident reporting systems.

To verify an incident, we first a geographic buffer around the extracted coordinates using Python libraries such as `shapely` and `geopandas`. This buffer filters incident records to those occurring within a specified proximity. For each candidate incident, we further check temporal consistency by comparing the incident's reported start time and duration with the tweet's timestamp. A successful match confirms the occurrence of an incident. In the absence of a matching incident, the tweet is classified as misinformation.

In summary, TrafficPulse offers a robust and scalable solution for filtering and verifying traffic incident reports from social media, thereby enhancing the reliability of integrated traffic prediction systems.

5. Experiments

In this section, we evaluate the performance of TrafficPulse—a pipeline designed to address key research questions central to our contributions. Specifically, we investigate: (i) *whether large language model-based techniques can reliably extract geospatial information from tweet-based incident reports*, and (ii) *whether our novel incident fact-checking approach can effectively detect and filter traffic-related misinformation.*

Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Classification Support Vector Machine (CSVM), and Classification Random Forest (CRF) serve as our baseline methods, following their successful application to traffic accident detection tasks in [16]. On similar traffic incident detection problems using sensor data, these methods have demonstrated strong performance, with CRF achieving the highest accuracy of 91.1% and precision of 92.0%. In contrast, CSVM achieved 88.5% accuracy with 88.3% precision.

Our experiments, conducted on a carefully curated benchmark, demonstrate a high extraction accuracy (90.29%) and strong misinformation detection performance, highlighted by a recall rate of 97.59%. These findings not only validate the integration of public data sources and advanced machine learning methods in traffic analysis but also underscore the potential of our approach for enhancing urban planning and public policy decision-making. Additional details about training and evaluation protocols can be found in Appendix D.

5.1. Benchmark Quality

We constructed a benchmark dataset comprising 175 carefully curated data points. Although the sample size is modest, these instances are highly valuable given the scarcity of traffic-related tweets within a raw pool of millions. The dataset is nearly balanced, with 83 positive and 92 negative samples, facilitating a reliable evaluation.

Moreover, the benchmark covers diverse traffic descriptions, with locations specified by roads—either intersections or segments between intersections and landmarks. This diversity in description styles poses a significant challenge for tasks that aim to extract precise location information.

5.2. Evaluation of the TrafficPulse Pipeline

To assess the performance of our proposed traffic misinformation detection pipeline, TrafficPulse, we conducted comprehensive experiments focusing on two critical components: the Traffic Information Extraction Module and the overall misinformation detection capability.

Traffic Information Extraction Module Performance. We evaluate research question (i) considering that accurate extraction of geo-coordinates is crucial for downstream analysis. We define extraction accuracy as:

$$Acc = \frac{\text{\# tweets with correctly extracted coordinates}}{\text{Total \# of tweets}}.$$

Our evaluation indicates an accuracy of 90.29% in identifying coordinates from tweets.

Traffic Misinformation Detection Performance. Similarly, we evaluate our misinformation detection pipeline through binary classification metrics, presented in Table 4. Comparing against established baselines from the literature including LDA, QDA, CSVM, and CRF [16], TrafficPulse demonstrates notably higher recall rates while maintaining competitive accuracy and precision. This superior recall performance is particularly crucial for misinformation detection, where identifying as many false reports as possible is a key priority. These results illustrate how TrafficPulse can not only extract traffic coordinates with usable accuracy but also effectively identify misinformation, advancing the state-of-the-art for social media-based traffic incident detection.

TABLE 4. PERFORMANCE METRICS DERIVED FROM THE CONFUSION MATRIX.

Model	Accuracy	Precision	Recall
LDA	85.8%	87.6%	45.2%
QDA	87.3%	89.4%	54.0%
CSVM	88.5%	88.3%	47.2%
CRF[16]	**91.1%**	**92.0%**	65.5%
TrafficPulse	90.29%	84.37%	**97.59%**

Limitations & Error Cases. Our analysis also revealed a challenging scenario in which TrafficPulse struggled to extract correct location coordinates from tweets mentioning multiple roads or highways. In such cases, the model failed to properly infer the spatial relationships between road segments (See examples in the Appendix C).

In this instance, TrafficPulse merely listed the road names without establishing the necessary connections between them. The tweet describes a road segment delineated by two distinct intersections, each defined by a pair of roads. This example highlights an important limitation of the current model, suggesting that further refinements are needed to accurately capture complex spatial relationships in route descriptions.

6. Conclusion

This paper introduces a method for integrating social media data with traffic prediction systems, addressing the challenges associated with tweet validation and enhancing the reliability of real-time transportation analysis. By leveraging large language models for geolocation extraction and proposing a method for incident fact-checking through public data sources, we establish a robust framework for filtering and validating traffic-related tweets. Furthermore, we introduce two new datasets linking traffic sensor data with tweet-based incident reports, providing valuable resources for future research in urban planning.

Our work paves the way for more accurate and trustworthy traffic models, highlighting the potential for combining social media with sensor data to improve traffic prediction, incident management, and urban planning. Moving forward, we aim to refine our validation pipeline and explore further applications of this integrated data for real-time decision-making and policy optimization in transportation science.

Acknowledgment

We gratefully acknowledge the support by the US Department of Transportation (USDOT) Tier-1 University Transportation Center (UTC) Transportation Cybersecurity Center for Advanced Research and Education (CYBER-CARE) (Grant No. 69A3552348332) and the Rice University Ken Kennedy Institute.

References

[1] J. Xiao, "Svm and knn ensemble learning for traffic incident detection," *Physica A: Statistical Mechanics and its Applications*, vol. 517, pp. 29–35, 2019.

[2] M. Yildirimoglu and N. Geroliminis, "Experienced travel time prediction for congested freeways," *Transportation Research Part B: Methodological*, vol. 53, pp. 45–63, 2013.

[3] Y. Tian, K. Zhang, J. Li, X. Lin, and B. Yang, "Lstm-based traffic flow prediction with missing data," *Neurocomputing*, vol. 318, pp. 297–305, 2018.

[4] A. Zhang, Q. Liu, and T. Zhang, "Spatial–temporal attention fusion for traffic speed prediction," *Soft Computing*, vol. 26, no. 2, pp. 695–707, 2022.

[5] W. Yao and S. Qian, "From twitter to traffic predictor: Next-day morning traffic prediction using social media data," *Transportation research part C: emerging technologies*, vol. 124, p. 102938, 2021.

[6] Y. Gu, Z. S. Qian, and F. Chen, "From Twitter to detector: Real-time traffic incident detection using social media data," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 321–342, Jun. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0968090X16000644

[7] Z. Zhang, Q. He, J. Gao, and M. Ni, "A deep learning approach for detecting traffic accidents from social media data," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 580–596, Jan. 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0968090X1730356X

[8] M. T. Zulfikar and Suharjito, "Detection Traffic Congestion Based on Twitter Data using Machine Learning," *Procedia Computer Science*, vol. 157, pp. 118–124, Jan. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S187705091931066X

[9] E. Alomari, R. Mehmood, and I. Katib, "Road Traffic Event Detection Using Twitter Data, Machine Learning, and Apache Spark," in *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2019, pp. 1888–1895. [Online]. Available: https://ieeexplore.ieee.org/document/9060213

[10] W. Yao and S. Qian, "From Twitter to traffic predictor: Next-day morning traffic prediction using social media data," *Transportation Research Part C: Emerging Technologies*, vol. 124, p. 102938, Mar. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0968090X20308354

[11] H. Chang, L. Li, J. Huang, Q. Zhang, and K.-S. Chin, "Tracking traffic congestion and accidents using social media data: A case study of Shanghai," *Accident Analysis & Prevention*, vol. 169, p. 106618, May 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0001457522000549

[12] E. M. Daly, F. Lecue, and V. Bicer, "Westland row why so slow?: fusing social media and linked data sources for understanding real-time traffic conditions," in *Proceedings of the 2013 international conference on Intelligent user interfaces*. Santa Monica California USA: ACM, Mar. 2013, pp. 203–212. [Online]. Available: https://dl.acm.org/doi/10.1145/2449396.2449423

[13] P. Panakkal and J. E. Padgett, "More eyes on the road: Sensing flooded roads by fusing real-time observations from public data sources," *Reliability Engineering & System Safety*, vol. 251, p. 110368, Nov. 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S095183202400440X

[14] D. A. Kurniawan, S. Wibirama, and N. A. Setiawan, "Real-time traffic classification with Twitter data mining," in *2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Oct. 2016, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/7863251/?arnumber=7863251

[15] C. Chen, "Freeway performance measurement system (pems)," Ph.D. dissertation, University of California, Berkeley, 2002.

[16] L. Zhu, B. Wang, Y. Yan, S. Guo, and G. Tian, "A novel traffic accident detection method with comprehensive traffic flow features extraction," *Signal, image and video processing*, vol. 17, no. 2, pp. 305–313, 2023.

Appendix

1. Keywords used for filtering

traffic, construction, commute, signal, crash, collision, vehicle, highway, congestion, northbound, southbound, truck, delay, accident, interstate, route, maintenance, closed, one-way, two-way, re- ported, off-ramp, toll, bridge, exit, incident, rd, st, ave, avenue, north, east, south, west, street, road, lane.

2. Prompts for Location Extraction

1) **Incident Detection:** The system queries, "Does the tweet explicitly report any incidents, accidents, or crashes in California?" Tweets failing this check are discarded.
2) **Road Mentions:** Next, it asks, "How many specific roads or freeways are mentioned in this tweet?" If no roads are mentioned, the tweet is filtered out.
3) **Single Road/Freeway Case:** The prompt, "Does this tweet include a precise street number (e.g., '9014 S Broadway, Los Angeles, CA')," is used to extract specific addresses when available. If a street number is not provided, the system checks for any mentioned landmarks to pinpoint the location.

4) **Two Roads/Freeways Case:** The system requests, "What are the roads or freeways mentioned? Please format the output as 'a and b, CA'." This allows us to interpret the intersection as a precise geolocation.
5) **Multiple Roads/Freeways Case:** For tweets mentioning more than two roads, the prompt is extended to, "What are the intersections of the roads/freeways mentioned? Format the output as 'a and b, CA; c and d, CA'." This enables extraction of multiple potential locations.

3. Location Extraction Failure Example

Tweet Example:*US 50 EB: from TOWER BRIDGE GTWY (West Sacramento) to Broadway (Sacramento).* https://t.co/wihGOS7eL8

TrafficPulse Extracted Location: *US 50 EB, TOWER BRIDGE GTWY and Broadway, CA*

True Location: *US 50 EB and Broadway, CA;*
US 50 EB and TOWER BRIDGE GTWY, CA

4. Training and Evaluation Details

Unlike traditional ML approaches, TrafficPulse's misinformation detection component does not require a train/test split because it follows a verify-by-reference approach rather than a trained classifier. The reported performance metrics (90.29% accuracy, 84.37% precision, 97.59% recall) are based on comparing the extracted coordinates from all benchmark tweets against the ground truth PeMS incident database using a geographic buffer and temporal matching algorithm. The baseline methods (LDA, QDA, CSVM, CRF) are results cited from other papers where those models were trained on sensor datasets, not implementations we created ourselves.

TrafficPulse uses ChatGPT-3.5-Turbo to extract location information via a structured decision tree of prompts, followed by geocoding through ArcGIS. Regarding the baseline ML models (LDA, QDA, CSVM, CRF), these are results cited from other papers where those models were trained on sensor datasets, not implementations we created ourselves. Our rule-based verification approach uses two main parameters: a geographic tolerance of 0.05 degrees to create a buffer around extracted coordinates and a ±10 minute temporal window to match incidents occurring around the tweet's timestamp.