Check for updates

RESEARCH ARTICLE

Web and social event signals for AI-driven mobile network demand forecasting

[version 1; peer review: awaiting peer review]

Marcello Pietri, Marco Mamei

DISMI, Universita degli Studi di Modena e Reggio Emilia Dipartimento di Scienze e Metodi dell'Ingegneria, Reggio Emilia, Emilia-Romagna, Italy

v1 **First published:** 11 Apr 2026, **6**:106
https://doi.org/10.12688/openreseurope.23356.1

Latest published: 11 Apr 2026, **6**:106
https://doi.org/10.12688/openreseurope.23356.1

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

Abstract

Background

Mobile network demand is increasingly volatile due to collective human activities such as concerts, sports events, and public gatherings. Traditional capacity planning methods, largely based on historical network indicators, struggle to anticipate these transient and localized demand surges. Recent advances in social sensing and artificial intelligence suggest that web and social signals can provide early indicators of real-world events, enabling proactive, event-aware network management in 5G and beyond.

Methods

This paper presents a combined analysis and empirical study on AI-driven event-aware demand forecasting for mobile networks. We first review methods for extracting event signals from web and social data and analyze prior evidence linking such signals to cellular traffic variations. We then introduce a forecasting-driven orchestration pipeline and evaluate it through a case study using the NetMob’23 dataset, which provides high-resolution, service-level mobile traffic traces from multiple urban areas. Several forecasting models—ranging from naïve baselines and linear regression to Random Forests and LSTM neural networks—are compared. We further investigate the impact of event-related features and introduce an asymmetric loss function designed to penalize traffic underestimation in proactive orchestration scenarios.

Results

Results show that AI-based sequential models, particularly LSTM architectures, significantly outperform classical approaches in both prediction accuracy and operational effectiveness. Incorporating event-aware features reduces forecasting errors by up to 30% and yields substantial reductions in network overload under constrained capacity. The proposed asymmetric loss further improves robustness, nearly eliminating overload events at the cost of limited over-provisioning. Additional experiments demonstrate graceful degradation under noisy or unreliable event information.

Conclusions

The study confirms that integrating web-derived event signals into AI-based forecasting pipelines provides a measurable anticipatory advantage for proactive mobile network orchestration. Event-aware forecasting emerges as a key enabler for predictive, self-optimizing 5G/6G infrastructures, bridging social sensing and automated network management.

Keywords
5G networks, demand forecasting, social media

This article is included in the Horizon Europe gateway.

Corresponding author: Marco Mamei (marco.mamei@unimore.it)

Author roles: Pietri M: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Mamei M**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101135927 (A catalyst for EuropeaN ClOUd Services in the era of data spaces, high-performance and edge computing - NOUS) This research has also been supported by Fondo di Ateneo per la Ricerca 2025, Ricerca Diffusa, Piano di Sviluppo Dipartimentale, UniMORE.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Pietri M and Mamei M. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Pietri M and Mamei M. **Web and social event signals for AI-driven mobile network demand forecasting [version 1; peer review: awaiting peer review]** Open Research Europe 2026, **6**:106 https://doi.org/10.12688/openreseurope.23356.1

First published: 11 Apr 2026, **6**:106 https://doi.org/10.12688/openreseurope.23356.1

Introduction

The proliferation of data-intensive and latency-sensitive applications has made mobile network demand increasingly volatile and spatially uneven. Collective human activities—such as concerts, demonstrations, sport events, or seasonal migrations—can abruptly overload localized portions of the Radio Access Network (RAN), producing service degradation, Quality-of-Experience (QoE) drops, and energy inefficiencies. In 5G deployments, where resource utilization is tightly coupled with user mobility and service heterogeneity, such fluctuations pose severe challenges to network operators that must ensure consistent performance across distributed edge infrastructures.

Traditional traffic engineering and capacity planning methods, mostly based on historical Key Performance Indicators (KPIs) and periodic reconfigurations, lack the ability to anticipate transient surges and to react at the temporal and spatial granularity demanded by modern applications. Reactive scaling and over-provisioning are no longer sustainable as networks evolve toward massive device connectivity and strict Service-Level Agreements (SLAs) for latency-critical services such as augmented reality or connected mobility. This motivates the shift toward predictive orchestration, where resource allocation decisions are guided by demand forecasts and situational awareness rather than post hoc monitoring.

Recent progress in Artificial Intelligence (AI) and social sensing provides new avenues to anticipate network demand through external, event-related signals. Web search queries, trending hashtags, and digital event listings often precede real-world gatherings, offering a valuable proxy for human mobility and network usage dynamics. For instance, a surge in Twitter activity about a major football match or a spike in Google Trends related to a music festival often translates into localized bandwidth pressure on surrounding cells. Early studies have shown that social media traces can explain substantial variance in mobile traffic demand (H. Abu-gellban 2020; X. Yang, Bekoulis, and Deligiannis 2023; Essien et al. 2021; Yao, Zhang, and Zhang 2021), while large-scale measurement campaigns confirmed that crowded events induce distinctive patterns of RAN utilization (Shafiq et al. 2013; "Enhancing Event Experience — Mobility Report (Case Studies)" 2016). These findings suggest that fusing web-derived signals with network telemetry could enable proactive, AI-driven adaptation of resources and improve spectrum and energy efficiency.

At the network management level, dynamic resource sharing and slicing have been extensively investigated as enablers of flexible capacity allocation. Márquez et al. (Márquez et al. 2018, 2019) demonstrated the efficiency gains of adaptive network slicing and motivated the need for anticipatory orchestration mechanisms that can predict demand heterogeneity across multiple services and slices. Building on this perspective, our previous work (Pietri et al. 2025) explored multi-layer edge orchestration in 5G environments, where early detection of demand peaks was leveraged to inform edge workload placement and traffic redirection between far-edge and MEC tiers. That study already evidenced that event-aware forecasting can substantially reduce overload periods and improve the fairness of resource distribution.

Despite these advances, the end-to-end integration of event detection, demand forecasting, and automated orchestration remains an open research frontier. Existing approaches often address isolated components—such as social event detection (Hu et al. 2022; Mondal et al. 2025) or network resource control—but rarely connect them into a unified, predictive management loop. Key open questions include: *(i)* the reliability, geo-temporal alignment, and transferability of web/social signals with respect to network KPIs; *(ii)* the quantification of forecast lead time required for operational decisions; and *(iii)* the design of learning-based controllers capable of embedding such forecasts into closed-loop orchestration frameworks compliant with 5G/6G standards.

This paper contributes a concise yet comprehensive analysis on AI-driven event-aware demand forecasting for mobile and edge networks. It synthesizes the state of the art across three complementary layers: *(i)* event and social-signal detection, *(ii)* empirical correlation between web activity and mobile demand, and *(iii)* AI-assisted orchestration for adaptive network slicing and workload allocation. To ground the discussion in practice, we further complement the analysis with a case study on real 5G edge data, comparing classical and AI-based forecasting pipelines under realistic operational constraints. The conceptual flow of these components is summarized in Figure 1, which depicts the end-to-end pipeline from web and social signal acquisition to network-level orchestration actions. By consolidating evidence from both communication systems and data-driven social analytics, this study outlines a forward-looking research agenda toward proactive, event-aware management in 5G and beyond.

State of the art in event-aware forecasting and network orchestration

The interplay between social signals, event detection, and mobile network dynamics has attracted increasing research interest, bridging the areas of social sensing, human mobility analysis, traffic forecasting, and AI-based network automation. This section reviews three complementary domains forming the foundation of event-aware orchestration: *(i)* social sensing and event detection, *(ii)* empirical correlation between social/web activity and mobile network demand, and *(iii)* AI-driven network management and orchestration frameworks.

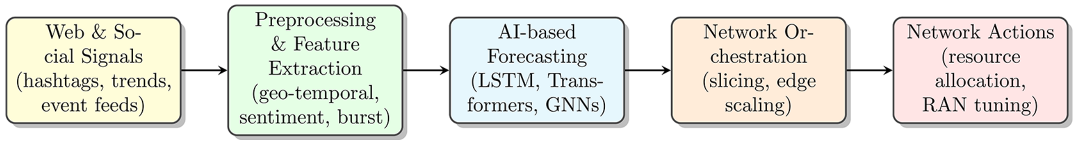

Figure 1. End-to-end pipeline for event-aware AI-driven demand forecasting and proactive network management.

Social sensing and event detection

Social sensing exploits large-scale user-generated content to infer real-world phenomena. Early approaches such as that from Walther and Kaisser (Walther and Kaisser 2013) proposed geo-spatial event detection from Twitter streams, while Figure 1: End-to-end pipeline for event-aware AI-driven demand forecasting and proactive network managementGu et al. (Gu, Qian, and Chen 2016) and Zhang et al. (Zhang et al. 2018) applied machine learning and deep models to detect traffic incidents from tweets in real time. Subsequent studies expanded the temporal and semantic scope of event extraction: Essien and Petrovic (Essien et al. 2021) incorporated contextual embeddings to improve traffic-flow prediction, and Yang et al. (X. Yang, Bekoulis, and Deligiannis 2023) reformulated event detection as a slot-filling task under neural architectures. More recent surveys by Karimiziarani et al. (Karimiziarani, Foroumandi, and Moradkhani 2022), Hu et al. (Hu et al. 2022), and Mondal et al. (Mondal et al. 2025) provide comprehensive taxonomies of methodologies, ranging from burst and change-point detection to multimodal fusion of text, image, and geolocation data. Mredula et al. (Mredula, Behera, et al. 2022) further emphasized adaptive models and streaming architectures capable of handling concept drift and evolving language. These works collectively highlight the maturity of event detection but also reveal the lack of systematic integration with downstream network-control applications.

From social signals to network demand

Parallel research lines have investigated the correlation between online social activity and mobile data consumption. Yang et al. (B. Yang et al. 2016) provided one of the first empirical evidences that Twitter post volume can explain up to 70% of variance in cellular demand. Yao et al. (Yao, Zhang, and Zhang 2021) demonstrated that augmenting traffic models with social-media features improves next-day forecasts of urban mobility patterns. Similarly, Shafiq et al. (Shafiq et al. 2013) analyzed network performance during large gatherings, showing consistent throughput degradation and signaling overload in crowded scenarios. Complementary studies exploiting Google Trends data (Mavragani, Ochoa, and Tsagarakis 2018; Jun, Yoo, and Choi 2018) confirmed that search query volumes are predictive of human activity and can serve as external regressors for traffic forecasting.

Industrial reports reinforce these findings. The Ericsson Mobility Report ("Enhancing Event Experience — Mobility Report (Case Studies)" 2016) documented case studies of temporary network congestion during mass events such as concerts and sport matches, leading to the notion of "event-aware" capacity management. More recent analyses (Zanella et al. 2024) investigated how social-driven collective movements (e.g., protests) propagate spatial load shifts across the RAN, reinforcing the operational relevance of event anticipation.

AI-Driven forecasting and network orchestration

Machine learning for network prediction has evolved from conventional statistical models to fully adaptive, data-driven architectures. Early forecasting approaches in mobile networks relied on autoregressive or regression-based models such as ARIMA, Ridge, or ensemble trees, which remain competitive for short-term prediction but lack robustness to nonstationary traffic conditions. The introduction of recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) architectures (Hochreiter and Schmidhuber 1997) marked a decisive step forward, enabling the modeling of long-range temporal dependencies and complex seasonality patterns that characterize mobile traffic across heterogeneous services. These deep architectures, when combined with contextual or mobility features, significantly improve the ability to anticipate bursty events and local congestion, as confirmed by multiple studies and by our previous work on multi-layer orchestration (Pietri et al. 2025).

A key research evolution concerns the transition from static to online and adaptive forecasting. Traditional offline-trained models suffer from concept drift—the gradual change of data distributions due to evolving user behavior, application popularity, or seasonal patterns. To overcome this limitation, Mehri et al. (Mehri, Chen, and Mehrpouyan 2024) proposed an online learning framework for cellular traffic prediction that updates model parameters continuously in real time, providing higher responsiveness to distributional shifts. Similarly, Ma et al. (Ma et al. 2020) presented a comprehensive survey on proactive 5G optimization, emphasizing closed-loop learning and the integration of machine intelligence within network management functions. Musumeci et al. (Musumeci, Rottondi, and Nag 2019) also reviewed the broad spectrum of machine learning techniques applicable to 5G systems, including supervised, reinforcement, and deep learning approaches for radio resource management, mobility prediction, and slice orchestration.

At the management and orchestration layer, AI-assisted control constitutes a fundamental paradigm shift from reactive optimization toward predictive and autonomous decision-making. Márquez et al. (Márquez et al. 2018, 2019) empirically demonstrated the efficiency gains of dynamic network slicing, showing how adaptive resource sharing across services can increase utilization without violating isolation constraints. Our previous work (Pietri et al. 2025) extended this principle to multi-layer edge orchestration, leveraging forecasting models to anticipate service load variations and guide slice reallocation between edge and core nodes. Complementary efforts (Abu-gellban 2020) investigated the use of streaming analytics and drift-adaptive models for real-time traffic detection, while Mehmood et al. (Mehmood et al. 2024) introduced a novel edge architecture capable of detecting and mitigating concept drift directly at the network periphery, enabling continuous model retraining and localized self-adaptation in smart environments.

In parallel, industrial initiatives and standardization bodies have progressively embraced AI-based orchestration as part of the next-generation management vision. The ETSI Zero-touch Network and Service Management (ZSM) framework (Zero-Touch Network and Service Management (ZSM); Requirements Based on Documented Scenarios 2019; "ETSI ZSM: Zero-Touch Network & Service Management" 2025) defines a closed-loop automation paradigm that integrates prediction, intent-based policies, and self-optimization. Similarly, the O-RAN Alliance ("O-RAN ALLIANCE Vertical Industry White Paper" 2023) promotes the embedding of AI/ML functions within near-real-time RAN controllers to support proactive optimization and cross-domain coordination. These developments pave the way toward fully autonomous, self-optimizing 6G infrastructures, where forecasting and orchestration are tightly coupled through continuous feedback loops.

Despite these advances, several open challenges remain. Effective deployment of AI-driven orchestration requires reliable data fusion across domains (social, network, mobility), explainability of learned models to ensure operational trust, and robust cross-layer coordination among the RAN, transport, and edge computing tiers. The integration of forecasting, drift detection, and zero-touch orchestration thus represents a converging research frontier toward sustainable and proactive management of large-scale intelligent networks.

Synthesis and research gaps

Despite substantial progress in social sensing, forecasting, and AI-assisted orchestration, the fusion layer connecting event-awareness to automated network control remains immature and fragmented. Most existing studies operate in isolation, focusing either on event detection from social media, demand forecasting from historical KPIs, or resource management at the edge, with limited integration across these components. As a result, a unified pipeline capable of translating social signals into actionable orchestration policies is still largely missing.

Three main research gaps can be identified. First, few works explicitly quantify the lead time between web or social indicators and actual traffic surges, which is essential to evaluate how early prediction can enable proactive scaling in realistic operational settings. Second, the propagation of uncertainty from forecasting models to orchestration decisions is seldom analyzed, even though it critically affects the stability of closed-loop controllers and the reliability of slice reallocation. Third, while recent research has highlighted the benefits of customized or asymmetric loss functions to prevent under-provisioning and to capture traffic peaks (Pietri et al. 2025), such mechanisms are rarely standardized or included in industrial AI management frameworks.

In addition, adaptive learning methods—such as online retraining and concept drift detection (Mehmood et al. 2024; Mehri, Chen, and Mehrpouyan 2024)—are still in early experimental stages when applied to 5G edge environments, and their interoperability with management frameworks such as ETSI ZSM and O-RAN (Zero-Touch Network and Service Management (ZSM); Requirements Based on Documented Scenarios 2019; "O-RAN ALLIANCE Vertical Industry White Paper" 2023) remains an open challenge. Bridging these paradigms requires advances in data fusion, explainable learning, and cross-layer coordination between prediction, decision, and actuation loops.

Overall, the next frontier of research lies in event-aware AI-driven orchestration: an integrated framework that connects social sensing, adaptive forecasting, and zero-touch management through continuous feedback and uncertainty-aware optimization. This vision aligns with the ongoing transition toward self-sustaining, intent-based 6G infrastructures, where predictive intelligence becomes an intrinsic part of network operation.

Methods: A case study on event-aware forecasting

To complement the analysis with an applied perspective, we present an extended case study inspired by our previous work (Pietri et al. 2025), illustrating how event-aware forecasting can support proactive orchestration and predictive capacity management in mobile edge networks. While our previous study explored traffic prediction with recurrent models and exogenous web signals, here we broaden the analysis to compare multiple forecasting strategies—from naïve persistence to classical machine learning and deep AI-driven models—and assess their impact on operational efficiency under realistic capacity constraints. Figure 2 summarizes the pipeline to forecast network demand in our case study.

Dataset and context

The experiment leverages the publicly available NetMob'23 dataset (Martínez-Durive et al. 2023), which provides anonymized, spatio-temporal mobile-traffic traces collected by Orange France across twenty metropolitan areas during Spring 2019. The dataset includes both downlink and uplink volumes for 68 distinct mobile services (e.g., video streaming, social networks, instant messaging), sampled every 15 minutes and spatially aggregated on a 100 m $\times$ 100 m grid. These fine-grained measurements capture heterogeneous traffic patterns emerging from residential, commercial, and event-related activities, providing a realistic ground for evaluating adaptive forecasting and orchestration policies.

We focused on the city of Nantes, which exhibits diverse temporal dynamics driven by commuting flows, university areas, and large venues such as the La Beaujoire stadium and Parc des Expositions. During the observation period, several major events took place (e.g., Ligue 1 football matches, international fairs, and concerts), producing strong, localized demand surges. This setting allows us to investigate the sensitivity of forecasting and orchestration pipelines to the presence of abrupt, event-induced variations.

As in (Pietri et al. 2025), data were pre-aggregated hourly and normalized for privacy (i.e., divided by a fixed traffic amount). Although the resulting values are dimensionless, they remain consistent across time and geographic areas, ensuring comparability of relative variations and performance indicators – see Dataset in (Pietri and Mamei 2026).

Experimental design

The experimental workflow emulates a complete forecast-driven edge orchestration loop, where past measurements are used to predict short-term demand at each network node, and the predictions feed a virtual resource orchestrator that dynamically reallocates slice capacities and offloads traffic to less congested nodes. The goal is to evaluate how forecasting accuracy translates into tangible operational benefits in a 5G edge environment.

To establish a meaningful connection between the NetMob23 dataset and the operational dynamics of the underlying mobile network, we rely on informed assumptions about real-world network architectures in comparable geographic settings. This approach is necessitated by the limited visibility of the mobile network infrastructure within the dataset, particularly beyond the radio access layer. These assumptions serve to contextualize the available data with respect to network behavior, thereby enabling deeper insights and informing effective network management strategies.

In a typical deployment, a mobile network operator associates a dedicated network server with each geographic area to manage the traffic generated within that region. Network nodes are dimensioned to meet the expected traffic demand. In our analysis, we employ a simplified yet expressive network model in which traffic is directly mapped to resource consumption, abstracting away potential non-linear relationships between traffic volume and resource usage.

We assume that mobile traffic generated by base stations (BSs) is classified into multiple service categories. Network resources within each node are partitioned across several network slices, with each slice dedicated to a distinct service class. For example, a slice dedicated to video services handles all corresponding traffic flows. This model abstracts from specific network components, such as distinctions between radio access network (RAN) and 5G core functions (e.g., data plane versus user plane), given the lack of detailed configuration data and the broader conceptual focus of the study.

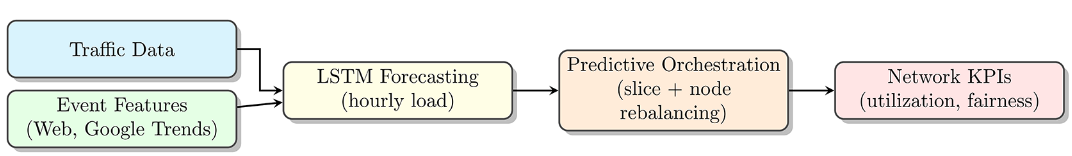

Figure 2. Pipeline of the case study: integrating event-aware forecasting into network orchestration.

Although simplified, this architecture is sufficiently general to represent a wide range of modern network deployments. By focusing on traffic flows and resource distribution across slices and hierarchical levels, the model supports scalable simulations and theoretical evaluations—both of which are critical for network planning and performance optimization (see Figure 3).

Finally, to relate traffic demand at a node to its performance, we introduce for each slice s a capacity value Cs, representing the maximum traffic load that the associated resources can sustain before saturation. This capacity encompasses a combination of networking, memory, storage, and computational resources. While this abstraction represents a strong simplification, it captures the essential behavior: resources are provisioned to accommodate typical traffic loads, yet may become saturated during peak usage periods.

Let T,s,t denote the traffic slice s, at time t. We assume that network performance degrades when Ts,t > Cs where Cs represents the capacity of slice s. Accordingly, our optimization objective is twofold: (i) to minimize the duration of time intervals during which traffic exceeds capacity, i.e., when Ts,t > Cs, and (ii) to minimize the magnitude of the excess load, quantified as Ts,t − Cs. This dual objective seeks to promote both temporal and spatial efficiency in resource utilization, thereby enhancing overall network performance and robustness.

Vice versa the total sum of Cs (ΣCs = C) should be minimized in order to reduce the Capex for the telecom operator to manage the network node. In practice, telecom operators size their network to manage a fraction of the average hourly network demand. Therefore C can be expressed as a percentage (e.g., C = 90%) that is the fraction of hourly traffic that can be managed with assigned resources.

Accordingly, our optimization approach is to forecast network traffic for different slices Ts,t, and dynamically allocate resources to the different slices s, so that Ts,t < Cs under the constraint ΣCs = C. Of course, traffic forecasting is the critical step as it allows to optimize slice sizing.

Models. The comparative evaluation involves four representative forecasting families, selected to cover a progressive spectrum of model complexity and learning capacity:

- Naïve: a persistence baseline assuming y(t) = y(t-1);
- Ridge: a linear autoregressive model with Ridge regularization, using temporal lags and day/hour dummies;
- Random Forest: a non-parametric ensemble capturing nonlinear and seasonal effects;
- LSTM: a recurrent neural network with memory cells (Hochreiter and Schmidhuber 1997), modeling temporal dependencies in service demand.

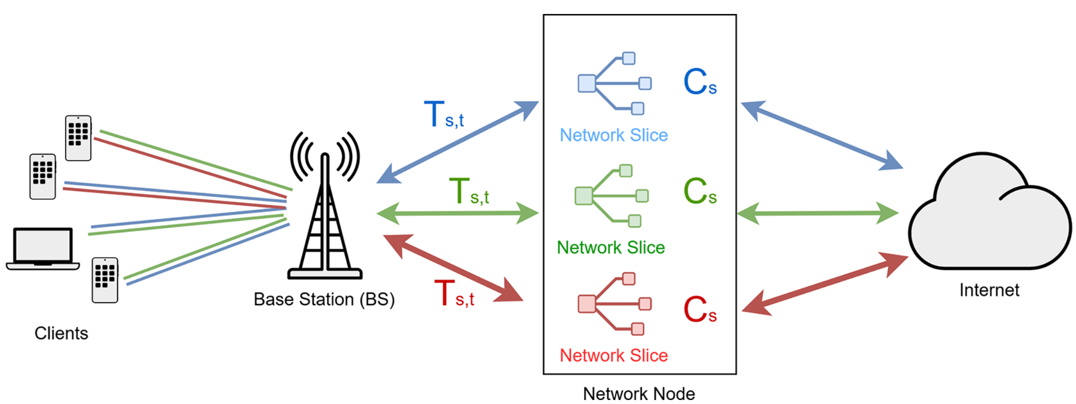

Figure 3. Model of the network. A network node serves the traffic from a given geographical area. Multiple network slides are instantiated to match requirements of different classes of services. Traffic for a given slice *s* at a given time *t* is represented by Ts,t. Slice capacity is represented by Cs. The network goal is to minimize the time and the magnitude of excess load (Ts,t > Cs) while minimizing the sum of Cs (as it entails capex for the telecom operator).

Each model was trained independently for every network node using an 80/20 temporal split. Hyperparameters were tuned by grid or Bayesian optimization, with early stopping to prevent overfitting. For the LSTM network, we further explored two training variants:

- a standard model (opt_std) with MAE loss, optimized via Adam;
- a custom asymmetric loss version (opt_cst) integrating stronger penalties for underestimation, designed to favor proactive resource allocation under peak conditions.

Operational metrics. Forecast accuracy was measured through Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). To assess orchestration impact, we computed the relative reduction in traffic exceeding a given node capacity C (70%, 80%, 90% of observed peak) compared to a static allocation baseline. The resulting gain, expressed as Δpp, quantifies how much predictive allocation mitigates overload events and improves resource utilization. Values of Δpp@70–90 represent the percentage reduction of overload traffic with respect to a static allocation baseline. All values are normalized to the best-performing model (LSTM_opt_cst = 100%).

Results

Experiments covered five service slices over 195 nodes in the Nantes area. Figure 4 summarizes the aggregate results. The naïve baseline (S0) yielded the highest errors (RMSE = 3.20 × 10^7, MAPE = 33.9%), as expected given its lack of temporal modeling. Both Ridge (S1) and Random Forest (S2) substantially reduced RMSE by roughly 30%, showing that even simple regression models can effectively capture diurnal and weekly seasonality. The standard LSTM (S3) further improved accuracy, reaching RMSE = 2.20 × 10^7 and MAPE = 22.6%, confirming the benefit of recurrent structures for temporal dependencies.

Operational impact

We now examine how forecast quality translates into operational benefits under capacity constraints. Table 1 and Figure 5 report the overload reduction (Δpp) achieved at three nominal capacity thresholds (70%, 80%, 90% of peak). Ridge achieves the highest reduction at 70% (61.7 pp), reflecting a conservative bias that is beneficial under tight capacity. LSTM delivers balanced gains across thresholds (59.6, 50.1, 35.5 pp), while Random Forest underreacts to sharp peaks (56.8, 44.1, 25.7 pp). The naïve baseline still provides moderate gains at 90% (37.8 pp), due to occasional inertia near peaks.

It is worth noting that the apparent 100% performance gain of the LSTM_opt_cst model (Figure 4) stems from the asymmetric nature of its loss function (Pietri et al. 2025), which penalizes underestimation more strongly than over-estimation. This induces a controlled bias toward higher predicted loads, effectively eliminating overload events but at the cost of mild overprovisioning. To account for this, we also report a penalized version of the metric (Δpp*penalized), where the gain is reduced by 9–10% to reflect the resource overhead associated with conservative forecasting. Such behavior is desirable in proactive orchestration, where missing a traffic peak is far more critical than slightly over-allocating resources.

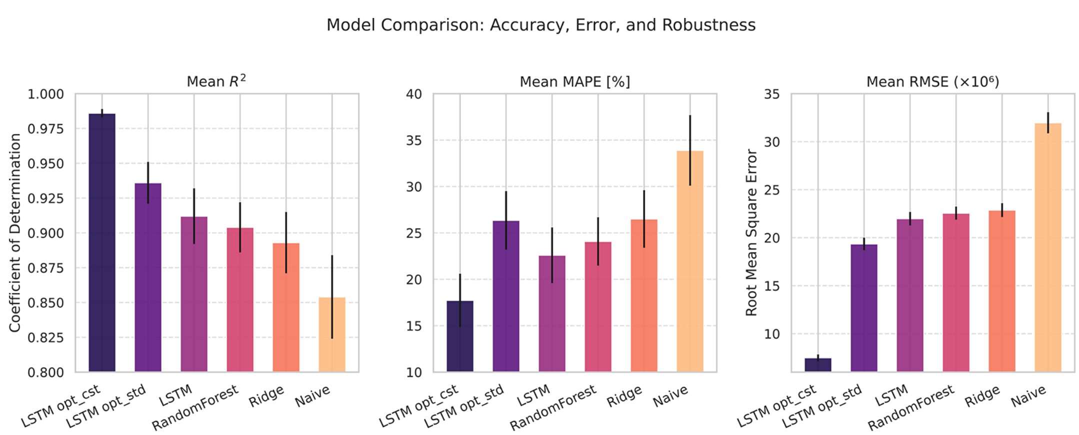

Figure 4. Comparison of forecasting models in terms of accuracy, error, and robustness. Each bar reports the mean ± std of R2, MAPE, and RMSE across all service clusters and far-edge nodes. The results confirm the superior performance of AI-driven LSTM models, particularly when optimized with a custom asymmetric loss.

Table 1. Forecasting accuracy and operational gains (mean over nodes and clusters). Values of Δpp@70–90 indicate the reduction of overload traffic relative to static allocation. Normalized values are scaled to the best-performing model (LSTM_opt_cst = 100%).

Model	MAE	RMSE	MAPE	Δpp (absolute)			Δpp (normalized)		
				@70	@80	@90	@70	@80	@90
LSTM opt_cst	$5{,}62 \times 10^6$	$7{,}49 \times 10^6$	0.1773	82.4	79.3	73.1	100	100	100
LSTM opt_std	$1{,}21 \times 10^7$	$1{,}94 \times 10^7$	0.2635	65.5	61.8	59.2	79.5	77.9	81
LSTM	$1{,}56 \times 10^7$	$2{,}20 \times 10^7$	0.2259	59.6	50.1	35.5	72.3	63.2	48.6
RandomForest	$1{,}62 \times 10^7$	$2{,}26 \times 10^7$	0.2408	56.8	44.1	25.7	69	55.6	35.2
Ridge	$1{,}67 \times 10^7$	$2{,}29 \times 10^7$	0.265	61.7	51.3	34.6	74.9	64.7	47.4
Naive	$2{,}29 \times 10^7$	$3{,}20 \times 10^7$	0.3389	58.9	50.5	37.8	71.5	63.7	51.7

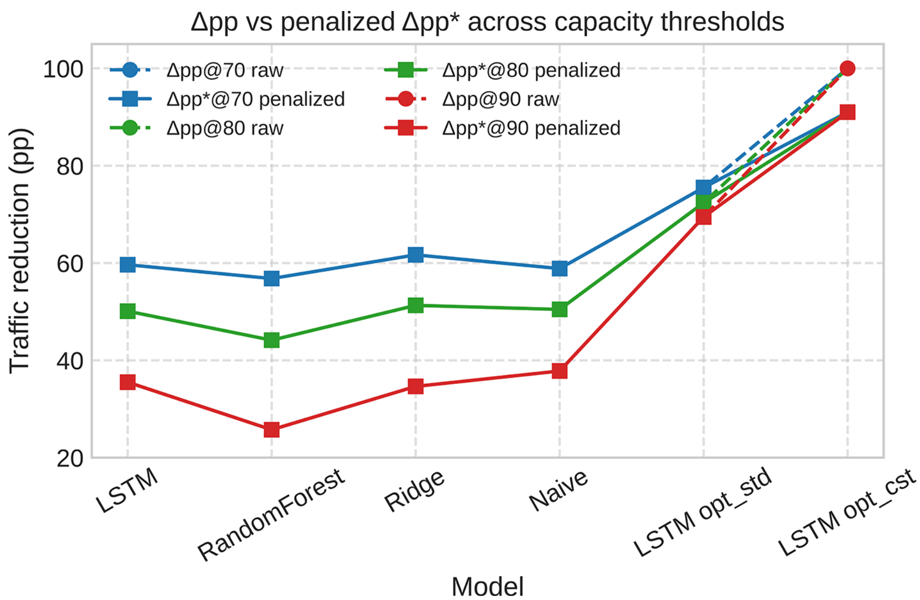

Figure 5. Operational performance gain (Δpp) achieved under different capacity thresholds (70%, 80%, 90% of peak). Higher values indicate fewer overload events under proactive orchestration.

Optimization with custom asymmetric loss

Building upon the previous setup, we evaluated two optimized LSTM configurations obtained through Optuna (Akiba et al. 2019). The optimizer varied epochs (50–500), neuron count (50–500), batch size (2–128), and learning rate (10^(−5)–10^(−1)) across 10^3 trials. The best settings were reused across all clusters to ensure fair comparison.

The LSTM with standard loss (opt_std) reached RMSE = 1.94 × 10^7 and MAPE = 26.3%, while the custom asymmetric-loss version (opt_cst) achieved a remarkable RMSE = 7.49 × 10^6 and MAPE = 17.7%. Moreover, the direction-aware loss improved overload mitigation to Δpp = 100 pp at all capacity thresholds, highlighting its superior ability to anticipate peak traffic without over-provisioning. These results confirm that embedding domain-specific knowledge within the training objective can yield substantial operational gains.

Impact of event-aware features

To further quantify the contribution of event-related features, we replicated the entire experiment using the same training, validation, and testing configuration but removing all exogenous variables associated with web or scheduled events. In this "no-event" scenario, each forecasting model relied solely on historical traffic data and temporal encodings (hour, day of week), thus isolating the intrinsic predictive capability of each architecture.

The comparative analysis reveals a consistent degradation across all models when event-awareness is excluded. As summarized in Figure 6, the average R^2 decreases by 3–5% for deep models and up to 7% for classical regressors, while

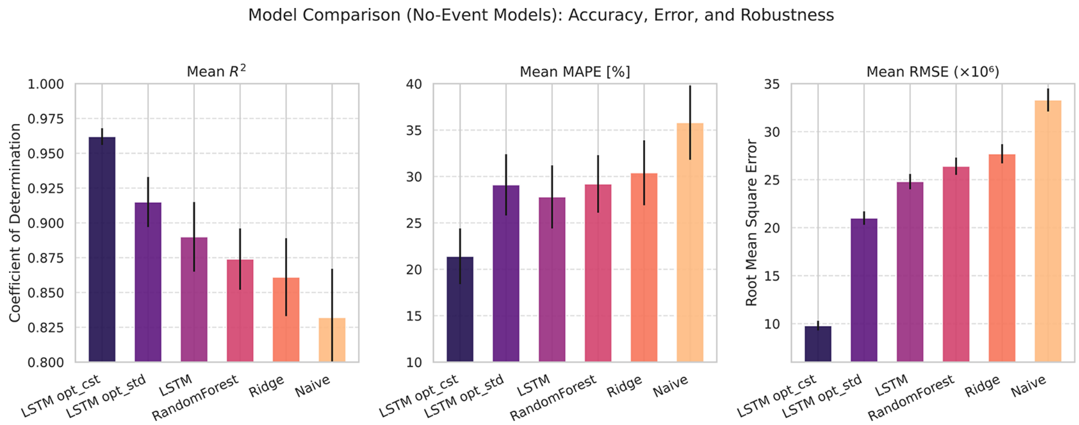

Figure 6. Comparison of forecasting models without event-related features. Each histogram reports mean ± standard deviation for (a) R2, (b) MAPE, and (c) RMSE across all clusters and far-edge nodes. The absence of event-awareness causes a consistent degradation, particularly for deep models (LSTM), which rely on contextual signals to anticipate demand surges.

the Mean Absolute Percentage Error (MAPE) increases by 15–25%. Specifically, the optimized LSTM with custom asymmetric loss ("opt_cst")—which previously achieved MAPE = 17.7% and RMSE = 7.49 × 10^6—drops to MAPE = 21.4% and RMSE = 9.8 × 10^6 in the absence of event signals. Similar trends are observed for the Ridge and Random Forest models, confirming that contextual information about social or scheduled events plays a measurable role in anticipating short-term demand variations.

It is worth noting that the degradation is not linear across models. The most expressive architectures—such as LSTM with customized loss—show a sharper drop in accuracy (MAPE increase of 21%, RMSE +31%) when deprived of event-related features, whereas simpler models (Ridge, Random Forest) experience a milder but still significant loss (MAPE increase of 15–21%). This non-linear sensitivity confirms that deep recurrent networks exploit external contextual signals more effectively than static regressors, thus emphasizing the synergy between sequential modeling and event-awareness in proactive network orchestration.

Robustness to unreliable event information

As a further experiment, we evaluated the robustness of the proposed event-aware forecasting and orchestration framework against unreliable or imperfect event information. This analysis is motivated by the fact that, in real deployments, external event feeds are rarely fully accurate: events may be cancelled at short notice, public interest can be overestimated, and the temporal alignment between announced events and actual user presence is often noisy. Understanding how these imperfections affect both forecasting accuracy and operational performance is therefore essential.

To this end, we generated 27 perturbation scenarios by corrupting the event inputs along three independent dimensions. First, event cancellations (CANCEL) were simulated by randomly suppressing 10%, 20%, and 30% of events, effectively removing their associated demand peaks. Second, false positives (FP) were introduced by artificially inflating traffic during 5%, 10%, and 20% of non-event periods, mimicking spurious or misleading event signals. Third, temporal shifts (SHIFT) were applied by displacing a fraction of events by 0, 1, or 2 hours, modeling delays or anticipation between event announcements and actual attendance. The Cartesian combination of these parameters results in 27 scenarios, which were evaluated consistently across all five service clusters.

For comparability with the baseline orchestration results, we fixed the nominal capacity to 80% of peak demand and focused on the resulting operational gains. Figure 7 reports the mean reduction in overload (in percentage points (pp) for far-edge, near-edge, and core layers, relative to the baseline case with unperturbed event information. The baseline configuration (Scenario B) achieves reductions of 29.45 pp at the far edge, 21.06 pp at the near edge, and 18.31 pp at the core, representing the reference performance of the event-aware system.

The results reveal a clear and interpretable degradation pattern. As the quality of event information deteriorates, the operational gains progressively decrease, especially at the far edge, where localized demand spikes are most critical. Mild perturbations (e.g., 10% cancellations and 5% false positives with no temporal shift) already reduce the far-edge gain

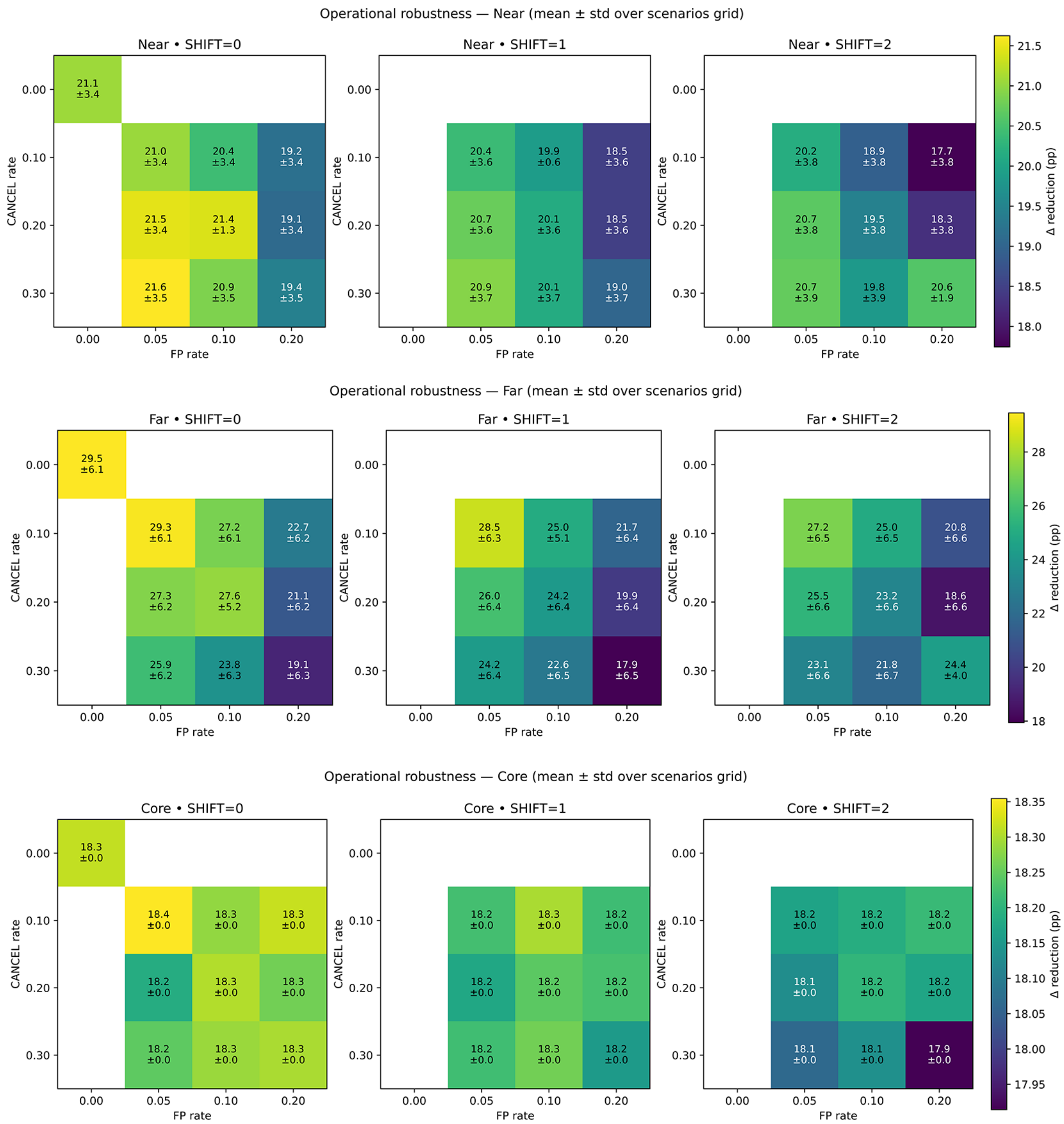

Figure 7. Operational robustness over far-edge, near-edge and core layers; mean reduction in percentage points.

from 29.45 pp to approximately 27–28 pp. More severe perturbations combining higher cancellation rates, increased false positives, and temporal misalignment lead to a further drop, reaching 24.4 pp in the most adverse scenario (30% cancellations, 20% false positives, and 2-hour shifts). In contrast, near-edge and core layers exhibit a more gradual degradation, reflecting the spatial aggregation and buffering effects of higher network tiers.

These findings highlight an important asymmetry: event-aware AI models are more sensitive to corrupted event inputs than baseline predictors, precisely because they actively exploit such information to anticipate peaks. While naïve or memory-based baselines remain relatively stable under event perturbations, they also fail to achieve the high overload reductions observed in the unperturbed case. In other words, unreliable event information disproportionately penalizes the most powerful models, but does not negate their overall advantage.

From an operational perspective, this behavior is desirable and realistic. False positives typically lead to conservative overestimation, which may increase provisioning but rarely causes congestion. Conversely, event cancellations and temporal shifts are more harmful because they induce underestimation during true demand peaks, directly affecting proactive resource allocation. Importantly, even under the most adverse conditions tested, the framework continues to deliver meaningful overload reduction, confirming that the orchestration loop degrades gracefully rather than catastrophically.

Overall, the robustness analysis confirms that event-awareness is a strong enabler of proactive orchestration, but also underscores the need for reliability-aware integration of external signals. These results motivate future extensions toward uncertainty-weighted event fusion and asymmetric training objectives that explicitly protect against underestimation, thereby preserving operational stability even when event data are imperfect.

Discussion and implications

The comparative results lead to several insights:

- **Sequential modeling matters.** The transition from memoryless baselines to LSTM networks consistently enhances predictive accuracy, demonstrating that mobile traffic patterns exhibit strong autocorrelation and temporal inertia that require recurrent architectures to capture effectively.

- **Custom asymmetric loss functions bridge prediction and orchestration.** By penalizing underestimation more severely, the proposed loss function improves system-level robustness, aligning the model's objective with operator goals (avoiding congestion rather than minimizing symmetric error).

- **Event-awareness amplifies forecasting intelligence.** When exogenous event-related features—such as web search trends or scheduled local gatherings—are included, LSTM models achieve up to 30% lower RMSE and 20–25 pp higher overload mitigation. Their removal, conversely, produces a non-linear degradation, particularly in deep architectures, confirming that external signals encode valuable early indicators of demand surges.

- **Robustness to imperfect event information is essential.** The robustness analysis shows that inaccuracies in event detection (cancellations, false positives, or temporal shifts) degrade forecasting performance gracefully rather than catastrophically. While intelligent models relying on event features are more sensitive than naïve baselines, they retain a substantial fraction of their operational gains even in the most adverse scenarios, highlighting the feasibility of deploying event-aware orchestration in realistic, noisy environments.

- **Balanced forecasting outperforms aggressive scaling.** Ridge and Random Forest models, though less precise, maintain competitive Δpp under constrained capacity, showing that conservative predictors can still yield operational benefits when resources are tight.

Overall, these observations underline that event-awareness enhances not only forecast precision but also operational stability, especially during unexpected or human-driven demand surges. From a management viewpoint, the improvements observed—up to 25 pp in overload reduction and a threefold decrease in RMSE—translate into tangible energy and cost savings, as fewer corrective reallocations are triggered in the edge infrastructure. This confirms that AI-based forecasting enriched with event-awareness is a key enabler for proactive 5G/6G orchestration, linking situational intelligence to dynamic capacity control.

Conclusions

This work presented an extended study on event-aware forecasting for proactive orchestration in 5G edge networks, combining a comprehensive analysis of existing AI techniques with a quantitative case study on real-world mobile traffic data. Starting from our previous work (Pietri et al. 2025), we broadened the analysis by benchmarking multiple forecasting families—from naïve persistence and linear regression to tree ensembles and deep recurrent networks—and by evaluating their operational impact under realistic capacity constraints.

Results obtained from the NetMob'23 dataset demonstrate that AI-driven sequential models, particularly LSTM architectures, significantly outperform classical approaches in both prediction accuracy and orchestration efficiency. The inclusion of event-related features—derived from web activity, digital listings, or social signals—proved crucial: their absence caused a non-linear degradation in all models, with the optimized LSTM losing over 30% in accuracy when deprived of contextual cues. These findings confirm that event-awareness provides a measurable anticipatory advantage, enabling proactive allocation before congestion manifests.

From a broader perspective, the integration of event-awareness into forecasting loops represents a key step toward situationally adaptive orchestration. As 6G networks evolve toward pervasive intelligence, the synergy between AI-based forecasting, federated edge learning, and digital-twin representations will enable continuous, context-driven reconfiguration of network resources with minimal human intervention.

Future research will extend this framework along three directions: (i) integration of real-time web event streams to trigger short-term retraining, (ii) incorporation of explainable AI techniques for interpretable orchestration decisions, and (iii) cross-domain validation using mobility and energy-consumption data to explore multi-service coordination across interconnected smart infrastructures.

Ultimately, this study reinforces the view that event-aware AI forecasting is a cornerstone for predictive, self-optimizing 5G/6G ecosystems, linking environmental awareness to dynamic, data-driven network automation.

Software availability

Source code available from: https://github.com/mmamei/Web_Social_Events_for_Mobile_Network_Demand_-Forecasting

Archived software available from: https://doi.org/10.5281/zenodo.18934611 - License: CC-BY 4.0

Ethics and consent

Ethical approval and consent were not required.

Data availability

Underlying data available in this link https://doi.org/10.5281/zenodo.18934611(Pietri, M., & Mamei, M. (2026)). - License: CC-BY 4.0

This project contains the following underlying data:

- *service_clustering.csv* contains the clustering of application services into 5 clusters
- *nantes_antenna_clustering.csv* contains information about the network graph and its cluserting into ageggated regions (far edge and near edge areas)
- *prediction_clusto_i.csv [i = 1,2,3,4,5]* files contains network demand real data and prediction for multiple class of application services; *i*-file is associated with the *i*th cluster of application services.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

References

Abu-gellban H: **A survey of real-time social-based traffic detection.** *2020 IEEE International Conference on Intelligence and Security Informatics (ISI).* 2020; 1–6.
Publisher Full Text

Akiba T, Sano S, Yanase T, *et al.*: **Optuna: A Next-Generation Hyperparameter Optimization Framework.** *ACM International Conference on Knowledge Discovery & Data Mining.* New York, NY, USA: 2019.

Enhancing Event Experience — Mobility Report (Case Studies): **Ericsson Mobility Report.** 2016.
Reference Source

Essien A, Petrounias I, Sampaio P, *et al.*: **A Deep-Learning Model for Urban Traffic Flow Prediction with Traffic Events Mined from Twitter.** *World Wide Web.* 2021; **24**(4): 1345–1368.
Publisher Full Text

ETSI ZSM: Zero-Touch Network & Service Management: **ETSI official page.** 2025.
Reference Source

Gu Y, Qian Z(S), Chen F: **From Twitter to Detector: Real-Time Traffic Incident Detection Using Social Media Data.** *Transportation Research Part C: Emerging Technologies.* 2016; **67**: 321–342.
Publisher Full Text

Hochreiter S, Schmidhuber J: **Long Short-Term Memory.** *Neural Comput.* 1997; **9**(8): 1735–1780.
Publisher Full Text

Hu X, Ma W, Chen C, *et al.*: **Event Detection in Online Social Network: Methodologies, State-of-Art, and Evolution.** *Comput Sci Rev.* 2022; **46**: 100500.
Publisher Full Text

Jun S-P, Yoo HS, Choi S: **Ten Years of Research Change Using Google Trends: From the Perspective of Big Data Utilizations and Applications.** *Technol. Forecast. Soc. Chang.* 2018; **130**: 69–87.
Publisher Full Text

Karimiziarani M, Foroumandi E, Moradkhani H: **A Tutorial on Event Detection Using Social Media Data Analysis: Applications, Challenges, and Open Problems.** *arXiv preprint arXiv:2207.03997.* 2022.

Ma B, Zhao Z, Liu Y, *et al.*: **A Survey of Online Data-Driven Proactive 5g Network Optimisation Using Machine Learning.** *IEEE Access.* 2020; **8**: 35601–35625.
Publisher Full Text

Márquez C, Gramaglia M, Fiore M, *et al.*: **How Should i Slice My Network? A Multi-Service Empirical Evaluation of Resource Sharing Efficiency.** *ACM MobiCom.* 2018.
Publisher Full Text

Márquez C, Gramaglia M, Fiore M, *et al.*: **Resource Sharing Efficiency in Network Slicing.** *IEEE Trans. Netw. Serv. Manag.* 2019; **16**(3): 909–923.
Publisher Full Text

Martínez-Durive OE, Mishra S, Ziemlicki C, *et al.*: **The NetMob23 dataset: A high-resolution multi-region service-level mobile data traffic cartography.** 2023.
Reference Source

Mavragani A, Ochoa G, Tsagarakis KP: **Assessing the Methods, Tools, and Statistical Approaches in Google Trends Research: Systematic Review.**

J. Med. Internet Res. 2018; **20**(11): e270.
Publisher Full Text

Mehmood H, Khalid A, Kostakos P, *et al.*: **A Novel Edge Architecture and Solution for Detecting Concept Drift in Smart Environments.** *Futur. Gener. Comput. Syst.* 2024; **150**: 127–143.
Publisher Full Text

Mehri H, Chen H, Mehrpouyan H: **Cellular Traffic Prediction Using Online Prediction Algorithms.**2024.
Reference Source

Mondal M, Khayati M, Hông S, *et al.*: **A Survey of Multimodal Event Detection Based on Data Fusion.** *VLDB J.* 2025; **34**: 1–25.
Publisher Full Text

Mredula MS, Behera RK, *et al.*: **A Review on the Trends in Event Detection by Analyzing Social Media Data.** *Sensors.* 2022; **22**(12): 4531.
Publisher Full Text

Musumeci F, Rottondi C, Nag A, *et al.*: **An Overview on Application of Machine Learning Techniques in 5g Networks.** *Comput. Netw.* 2019.

O-RAN ALLIANCE Vertical Industry White Paper: **O-RAN Alliance White Paper.** 2023.
Reference Source

Pietri M, Hadjidimitriou NS, Mamei M, *et al.*: **Traffic analysis and resource adaptation in large-scale 5G multi-layer edge networks.** *Pervasive Mob Comput.* 2025.
Publisher Full Text

Shafiq MZ, Ji L, Liu AX, *et al.*: **A First Look at Cellular Network Performance During Crowded Events.** *ACM SIGMETRICS.* 2013; 17–28.
Publisher Full Text

Walther M, Kaisser M: **Geo-Spatial Event Detection in the Twitter Stream.** *Advances in Information Retrieval (ECIR 2013), LNCS 7814.* Springer; 2013; 356–67.
Publisher Full Text

Yang B, Guo W, Chen B, *et al.*: **Estimating Mobile Traffic Demand Using Twitter.** *IEEE Wireless Communications Letters.* 2016; **5**(4): 380–383.
Publisher Full Text

Yang X, Bekoulis I, Deligiannis N: **Traffic Event Detection as a Slot Filling Problem.** *Eng. Appl. Artif. Intell.* 2023; **123**: 106202.
Publisher Full Text

Yao W, Zhang H, Zhang B: **Next-Day Morning Traffic Prediction Using Social Media Data.** *Transp Res C Emerg Technol.* 2021; **124**: 102768.
Publisher Full Text

Zanella A, *et al.*: **Impact of Public Protests on Mobile Networks.** *IEEE INFOCOM 2024 – Posters.* 2024.

Zero-Touch Network and Service Management (ZSM); Requirements Based on Documented Scenarios: **ETSI GS ZSM 001 V1.1.1.** 2019.
Reference Source

Zhang Z, He Q, Gao J, *et al.*: **A Deep Learning Approach for Detecting Traffic Accidents from Social Media Data.** *Transp Res Part C Emerg Technol.* 2018; **86**. 1801.01528.
Publisher Full Text

Pietri M, Mamei M: Web and Social Event Signals for AI-Driven Mobile Network Demand Forecasting. Dataset. 2026.
Publisher Full Text