LOCATION-AWARE MACHINE LEARNING TO FORECAST URBAN FOOT TRAFFIC

DOI
https://doi.org/
10.18690/um.fov.4.2026.38

ISBN
978-961-299-152-4

KATAYOON POURMAHDI, BAHAREH NASERI,
ILIA GUGENISHVILI, JOZSEF MEZEI,
ANNA-GRETA NYSTRÖM

Åbo Akademi University/Faculty of Social Sciences, Business and Economics, Turku, Finland
katayoon.pourmahdi@abo.fi, bahareh.naseri@abo.fi, ilia.gugenishvili@abo.fi, jozsef.mezei@abo.fi, anna-greta.nystrom@abo.fi

Smart cities increasingly rely on intelligent tools to manage urban mobility, reduce congestion, and support evidence-based public governance. This study examines how machine learning (ML) models can support urban mobility management by predicting footfall and multimodal traffic flows. Five ML models were trained and evaluated on a dataset of 261,539 hourly observations collected from a 16-sensor monitoring network across four urban locations in Turku, Finland. The modelling framework integrates temporal, meteorological, socio-economic, and spatial covariates to capture location-specific mobility flows. To prevent temporal leakage, model evaluation used a strict chronological train–test split, supported by five-fold time-series cross-validation; all metrics are reported with empirical standard deviations (±SD). XGBoost recorded the strongest predictive performance (R^2 = 0.720 ±0.071 in the City Centre, 0.839 ±0.099 in the University corridor, and 0.570 ±0.105 in the cultural district). Location 4, a low-volume remote sensor, yielded R^2 = 0.179 ±0.231, reflecting the inherent predictability limits of a near-zero-volume sensor (avg. 2.4 pedestrians/h). The findings guide sensor deployment planning, real-time mobility management, and data infrastructure design in smart city contexts.

Keywords:
smart city,
footfall flows,
mobility management,
machine learning,
neural networks

University of Maribor Press

1 Introduction

Urban governance and management practices evolve rapidly under the influence of various factors, such as climate change, migration, and digitalization (Peráček & Kaššaj, 2026). In the digital age, urban management involves working with large volumes of urban data and requires coordination among authorities, stakeholders, and citizens to co-design and co-produce solutions. The situation places the emphasis on the ability to translate data into actionable insights for effective decision-making (Gil et al., 2026). Technological tools and practices, such as sensors, real-time monitoring, and spatial data analytics, are important in that process (Jamal et al., 2026; Lingani et al., 2019).

Policymakers seeking to improve urban governance and decision-making have adopted machine learning (ML) methods (Jamal et al., 2026) to forecast traffic flow and pedestrian movement patterns (Haghshenas et al., 2026; Koech & Sobhani, 2026). Despite the recent growth in the use of urban data infrastructure, the application of ML to support urban governance remains under-researched across different contexts (Lingani et al., 2019; Wolniak et al., 2024). This gap is particularly evident in the area of urban mobility, which encompasses both pedestrian movements and multimodal traffic. The topic is one of the most complex and resource-intensive facing city administrations. Much of the existing research evaluates the technical performance of ML methods for traffic prediction and overall predictive accuracy (Kim & Chan, 2025; Wu et al., 2024). It typically reports global performance metrics aggregated across cities or sensor networks, with limited attention to model variability across different spatial contexts. As a result, we know relatively little about how and why ML model performance changes across heterogeneous urban contexts and its implications for the design of context-aware mobility prediction systems. An additional research gap concerns the limited translation of predictive outputs, such as footfall forecasts, congestion risks, or short-term demand surges, into actionable insights for urban planners, traffic managers, and public administrators. Although ML methods can generate accurate predictions, few studies connect these outputs to concrete governance applications or real-time operational decision-making.

Concurrently, the consequences of insufficient mobility management are wide-ranging: traffic congestion driven by inefficient traffic control, excess demand, population growth, or weather disruptions continues to degrade air quality, extend travel times, and diminish the quality of life for city residents (Lingani et al., 2019). In contrast, accurate and context-sensitive predictions of footfall and multimodal traffic flows offer opportunities to identify emerging congestion hotspots, optimize traffic control, and support responsive, sustainability-oriented urban mobility strategies.

This study aims to predict footfall and multimodal traffic flows to support effective urban management and decision-making, thereby contributing to safer streets, reduced pollution, and a more efficient use of urban infrastructure. To do so, we apply five ML models to predict day-to-day traffic trends using mobility data alongside meteorological and socio-economic datasets. We report location-specific model performance and identify the key factors driving predictions, with the goal of producing findings that are directly useful to urban planners and policymakers. The following research questions guide this study:

RQ1: In which urban contexts, and with what performance variability, can ML models reliably predict pedestrian footfall and traffic flows?
RQ2: Which contextual factors are the most influential predictors of urban pedestrian and traffic flows, and what do interpretable ML techniques reveal about their relative contributions?

To address these questions, we draw on a multi-sensor dataset of 261,539 hourly observations collected across 16 monitoring devices in Turku, Finland. Predictor variables include short-term temporal lag features, cyclical time-of-day encodings, meteorological conditions (temperature and precipitation), and socio-economic and spatial covariates (household income and distance from the City Centre).

2 Literature Review

ML has broadened the decision-support options available to mobility managers in smart cities. They can now access large volumes of real-time traffic data, road network information, and traveller behaviour patterns (Kim & Chan, 2025). Consequently, city managers can dynamically reallocate traffic flows, reduce congestion, and enhance overall mobility efficiency (Alhasnawi et al., 2025). A growing body of research recognizes that potential and explores predictive models underpinning smart mobility applications. A smaller (but growing) research strand expressly addresses the science of location-specific or context-stratified mobility prediction, indicating that model performance varies substantially across urban settings (Schläpfer et al., 2021). Comparative studies assessing multiple classification and regression techniques, including deep neural networks (DNNs), support vector machines (SVMs), decision trees, random forests (ensembles of decision trees aggregated by majority vote or averaging), and discrete choice models (Ali, 2025; Fernández-Delgado et al., 2014), establish that SVMs and random forests are consistently among the top-performing classifiers. For example, Fernández-Delgado et al. (2014) demonstrate that tree ensembles and kernel methods rival or outperform deep networks across a broad range of classification tasks. That learning guided the choice of XGBoost (a regularized gradient-boosted tree ensemble) over random forest: XGBoost incorporates L1/L2 regularization and is computationally more efficient on tabular data (Chen & Guestrin, 2016), making it better suited to the sensor-level modelling context. These findings underscore the advantages of non-linear models for capturing the complexity of urban mobility behaviour. Traditional statistical methods, such as logistic and linear regression, often fail to reproduce these dynamics with comparable accuracy (Dabiri & Heaslip, 2018; Fernández-Delgado et al., 2014; Pulugurta et al., 2013).

Recent studies signal the relevance of integrating spatial and temporal relationships in predictive models for footfall and traffic flows, and deep learning for footfall and traffic flow prediction (cf. Asher et al., 2023; Ng et al., 2017). For instance, Koech and Sobhani (2026) and Murcio and Wang (2025) demonstrate variations in pedestrian activity across several locations in Dublin and the UK, emphasizing the need for context-aware modelling. Similarly, Wang et al. (2025) showed that including weather variables alongside temporal features improves short-term traffic flow predictions, with their hybrid CNN–BiLSTM model outperforming five baseline models.

The above studies establish that ML model performance is sensitive to algorithmic design and also to the spatial and socio-economic characteristics of the data environment. Such sensitivity has direct implications for how predictive outputs, particularly footfall estimates, should be interpreted and operationalized in practice in urban locations. That also underscores the need for models that account for contextual variability to support reliable, actionable decision-making in smart mobility contexts and systems.

3 Modelling

3.1 Data Preprocessing

The data used in this study were gathered in Turku, Finland, from a network of 16 pedestrian and traffic monitoring devices (referred to as device01–device16) deployed across four spatially distinct urban locations. The raw dataset comprises 261,539 hourly observations recorded continuously between March 31 and November 1, 2025. Each data point corresponds to a specific device, measurement scenario (defined here as a unique directional or aggregated counting mode of a sensor, e.g., inbound, outbound, or total flow, as recorded by the device firmware), date, and hour, and includes counts of six traffic categories: pedestrians (Human), cyclists (Bike), cars (Car), buses (Bus), trucks (Truck), and other vehicles (Other Vehicle). Across the full dataset, pedestrian counts exhibit the widest variation (mean = 172.1, SD = 287.2, range: 0–11,466), reflecting strong spatial and temporal heterogeneity in foot-traffic intensity. Car counts (mean = 23.4, SD = 124.2) and bicycle counts (mean = 8.9, SD = 24.6) show comparatively lower but still substantial variance, while counts for buses, trucks, and other vehicles are sparse.

In addition to traffic counts, each record includes contextual covariates: hourly air temperature (range: −6.1 to 29.9 °C) and precipitation (range: 0–10.6 mm) sourced from Finnish Meteorological Institute (FMI) data, the average household income of the surrounding postcode (range: €29,878–€42,046)[1], and the geodesic distance of each device from the City Centre (range: 0.05–1,529 m). The 16 monitoring devices generated sixty-four directional and aggregated scenarios, enabling both direction-specific and total-flow analyses. For modelling purposes, the devices were grouped into four locations based on geographic proximity and urban function: Location 1 (City Centre / Tori, devices01 and 03–12, n = 128,194 records after preprocessing), Location 2 (Hämeenkatu / University corridor, device02, n = 10,259), Location 3

[1] Average rather than median income is reported here because Statistics Finland publishes household income at the postal-code level as an arithmetic mean. We acknowledge that right-skewed income distributions may cause this proxy to overweight high earners.

(Logomo / Ratapihankatu cultural district, devices13–15, n = 25,337), and Location 4 (remote building entrance, device16, n = 2,263).

Before model training, temporal lag features (lags 1, 2, and 3 of pedestrian, car, bus, truck, and bicycle counts) and cyclical hour encodings (sine and cosine transformations of the hour-of-day variable) were appended to each record. Missing values were present in two forms: (i) structurally missing lag values, arising inevitably from lag construction at the beginning of each device's time series (the first three records per device lack valid lag-1, lag-2, and lag-3 observations); and (ii) sporadic missing values in the temperature variable, attributable to sensor recording gaps, which were imputed using the column median before lag construction. All remaining records with any missing value were removed via listwise deletion. This preprocessing yielded a final analytical dataset of 166,053 complete observations (a retention rate of 63.5% of the raw 261,539 records), which was used for all model training and evaluation.

3.2 Input Features

A total of twenty-eight input features were constructed and used for model training. They were organized into five conceptually distinct groups: temporal, lag-based, meteorological, socio-economic, and spatial. Table 1 provides a full list of feature names, group membership, and a brief description of each feature.

3.3 Model Development

We addressed the two research questions by selecting five ML models to span a range of algorithmic complexity and interpretability. The selection includes a parametric linear baseline (linear regression), a shallow non-linear tree-based model (Decision Tree), a powerful gradient-boosting ensemble (XGBoost), a fully connected artificial neural network (ANN), and a recurrent sequential architecture (LSTM). The following subsections describe the architecture and training configuration of each model in detail.

Linear Regression: parametric baseline assuming a linear relationship between features and traffic counts; used to quantify the non-linearity gain of more complex models.

Table 1: Complete list of input features used for model training (n = 28)

Feature Groups	Feature Names	n	Description
Temporal (cyclical)	Hour_sin, Hour_cos	2	Sine and cosine transformations of the hour-of-day variable (0–23) to encode circular time continuity.
Temporal (categorical)	Weekdays	7	Binary dummy variables for each day of the week (one-hot encoded, 7 columns).
Pedestrian	Human_lag1, Human_lag2, Human_lag3	3	Pedestrian counts at t−1, t−2, and t−3 hours capture short-term autocorrelation in foot traffic.
Car	Car_lag1, Car_lag2, Car_lag3	3	Car counts at t−1, t−2, and t−3; top-ranked predictors in feature importance analysis.
Bus	Bus_lag1, Bus_lag2, Bus_lag3	3	Bus counts at t−1, t−2, and t−3 reflect the co-occurrence of public transport with pedestrian flows.
Truck	Truck_lag1, Truck_lag2, Truck_lag3	3	Truck counts at t−1, t−2, and t−3; second-highest importance group after car lags.
Bicycle	Bike_lag1, Bike_lag2, Bike_lag3	3	Bicycle counts at t−1, t−2, and t−3 capture active mobility co-movement with pedestrians.
Meteorological	Average temperature [°C], Precipitation [mm]	2	Hourly air temperature (−6.1 to 29.9°C) and precipitation (0–10.6 mm) from the Finnish Meteorological Institute.
Socio-economic	Average income of households	1	Average annual household income of the device's postal-code area (€29,878–€42,046).
Spatial	Distance from the City Centre	1	Geodesic distance of each monitoring device from the Turku City Centre (0.05–1,529 m).

Decision Tree: non-parametric, rule-based model (maximum depth 6) providing interpretable predictions; serves as a transparent non-linear baseline complementing the parametric linear regression baseline.

XGBoost: gradient-boosted decision tree ensemble (200 estimators, depth 4, learning rate 0.05) with L1/L2 regularization; one of the best-performing models for structured tabular data. XGBoost was selected over random forest, which also ranks among the top-performing models (Fernández-Delgado et al., 2014), because it offers equivalent or superior predictive performance on tabular sensor data while providing built-in L1/L2 regularization, native handling of missing values, and gain-based SHapley Additive exPlanations (SHAP) feature importance (Doshi-Velez & Kim, 2017).

ANN: fully connected feedforward network with three hidden layers (256–128–64 neurons), batch normalization, ReLU activation, and dropout; trained over 150 epochs using the Adam optimiser.

LSTM: A two-layer recurrent network (128 hidden units, sequence length of 6 hours, dropout rate of 0.2) designed to capture temporal dependencies in sequential traffic data; trained on a 30,000-sequence subsample over 100 epochs.

All models were evaluated on a chronological 80/20 train-test split, with performance reported as mean ±SD across 5-fold time-series cross-validation. A LAG_ONLY baseline (XGBoost on $t-1$ lags only) confirmed that full models learn beyond autocorrelation. Input features were standardized using a StandardScaler fitted on the training set alone.

4 Findings

Under the revised temporal split (chronological 80/20 train–test), XGBoost delivered the strongest location-stratified performance across all four monitoring zones (best: Location 2, $R^2 = 0.839 \pm 0.099$; Location 1, $R^2 = 0.720 \pm 0.071$). The ANN performed comparably (Location 2 $R^2 = 0.832 \pm 0.103$; Location 1 $R^2 = 0.717 \pm 0.061$). Parametric baselines were weaker (LR: $R^2 = 0.511$–0.514; DT: $R^2 = 0.537$–0.645 at Locations 1–2). The LAG_ONLY autocorrelation baseline (XGBoost using $t-1$ features only) yielded $R^2 = 0.141$–0.425, confirming substantial contextual gains (Location 1: $\Delta R^2 = +0.295$; Location 2: $\Delta R^2 = +0.698$; Location 3: $\Delta R^2 = +0.389$; Location 4: $\Delta R^2 = +0.290$) beyond simple autocorrelation.

When models were trained separately for each of the four monitoring locations, prediction accuracy varied, revealing strong spatial heterogeneity in predictability. Location 2 (Hämeenkatu / University corridor) yielded the highest accuracy across all models, with XGBoost achieving $R^2 = 0.839$ (± 0.099) and RMSE = 118.9. This location exhibits a highly regular noon-hour pedestrian peak ($\approx$ 600 persons/h at 12:00), consistent with lunchtime activity at the University corridor, which ML algorithms can capture effectively. In contrast, Location 3 (Logomo / Ratapihankatu cultural venue) produced the weakest results (XGBoost $R^2 = 0.570 \pm 0.105$), reflecting the irregular, event-driven footfall patterns typical of arts and entertainment districts. Location 4 (Remote / Outlier, device 16, avg. 2.4

pedestrians/h) yielded the lowest predictive accuracy across all models (XGBoost $R^2 = 0.179 \pm 0.231$, RMSE = 1.46). The near-zero pedestrian counts at this remote building entrance produce a degenerate series with minimal signal for ML models to exploit. Notably, the LAG_ONLY baseline also failed at this location ($R^2 = -0.111$), confirming that Location 4's low predictability is intrinsic to the measurement context rather than a limitation of the modelling approach.

All locations exhibit strong diurnal rhythms, with pedestrian activity concentrated during the daytime and approaching zero in the early hours of the morning. Location 1 (City Centre / Tori) shows the broadest activity window, reflecting the intensity of the commercial core throughout the day. Location 2 shows the sharpest single peak (12:00, ≈ 600 persons/h), consistent with midday university activity. Location 3 displays a broader, flatter profile with a secondary afternoon peak around 14:00 (≈ 208 persons/h). These contrasting temporal signatures confirm that time-of-day encoding, and lag features are essential inputs for modelling.

Table 2 summarises the full performance results. XGBoost feature importance (gain-based) and SHAP beeswarm analysis (Location 1) are provided as supplementary figures. Across Locations 1–3, the dominant predictor was Hour_cos (mean gain = 0.277), followed by Human_lag1 (0.117) and average household income (0.108). Location-specific rankings reveal an important divergence: at Locations 1 and 2, temporal features dominate (Hour_cos gains of 0.288 and 0.295, respectively), whereas at Location 3 (Logomo), household income ranked first (gain = 0.325), substantially above Hour_cos (0.248), reflecting the socio-economic stratification of the arts and entertainment district catchment. Distance from the City Centre was a consistent spatial predictor across all main locations (mean gain = 0.031).

5 Discussion

With regard to RQ1, XGBoost and ANN emerged as the most accurate models for hourly foot-traffic and consistently outperformed parametric baselines. In the location-stratified evaluation using a strict temporal split, XGBoost achieved the strongest predictive performance ($R^2 = 0.720 \pm 0.071$ in the City Centre, 0.839 ±0.099 in the University corridor, and 0.570 ±0.105 in the cultural district), consistent with broader evidence that ensemble methods outperform parametric baselines for urban traffic forecasting (Wang et al., 2025). In contrast, the LSTM model underperformed relative to its theoretical potential, primarily due to training on a limited subsample and early overfitting; a limitation also noted by Dabiri and Heaslip (2018), who highlight that deep sequential models require substantial training data to realize their full capacity. The slight advantage of ANN at the high-volume City Centre location suggests that neural networks scale more effectively with larger and more heterogeneous training sets (Ali, 2025), whereas XGBoost is preferable when data are limited, or interpretability is required (Doshi-Velez & Kim, 2017).

With regard to RQ2, XGBoost feature importance and SHAP analysis reveal a consistent predictor hierarchy. The cosine encoding of hour-of-day (Hour_cos) was the dominant predictor across all main locations (gain = 0.288–0.295), capturing the strong diurnal rhythm of pedestrian flows. Human_lag1 ranked second at Locations 1 and 2 (gain = 0.122–0.201), reflecting short-term autocorrelation. A noteworthy finding is that at Location 3 (Logomo), average household income ranked first (gain = 0.325), above Hour_cos (0.248), reflecting the socio-economic stratification of the arts district catchment. Distance from the City Centre was a consistent spatial predictor (mean gain = 0.031), while meteorological variables exerted secondary effects. The SHAP beeswarm plot (Location 1) confirms that high Hour_cos values drive strongly positive contributions (SHAP ≥ +100), and that distance from the City Centre shows negative contributions at large distances, consistent with declining pedestrian density away from the urban core. This finding is consistent with those of Koech and Sobhani (2026) and Murcio and Wang (2025) showing that temporal and spatial variables are significant in urban movement prediction.

Prediction performance is driven more by urban land-use type than by data volume. The University corridor achieved the highest accuracy despite having the fewest training records (n = 10,259), because its pedestrian pattern was highly regular and institutionally structured. In contrast, the Logomo cultural venue produced the weakest results (R^2 = 0.570 ±0.105) due to irregular, event-driven footfall; incorporating event calendar data at such locations is a clear direction for future improvement. Location 4 (remote sensor, avg. 2.4 pedestrians/h) yielded R^2 = 0.179 ±0.231; the LAG_ONLY baseline also failed here (R^2 = −0.111), confirming that low predictability at this site is intrinsic to the measurement context rather than a modelling limitation. This establishes a case for a tiered strategy: location-specific models for atypical sensors, and shared models for high-volume locations.

Table 2: Prediction performance of ML models by location

Locations	Model	R^2 (±SD)	RMSE (±SD)	MAE (±SD)	MSE	MAPE (%)
Location 1: City Centre / Tori: devices 01, 03–12; market square area; avg. distance from City Centre = 69 m; n = 128,193 records.	ANN	0.717 ±0.061	109.34 ±22.03	44.38 ±6.45	11,955	142.9%
	XGB	**0.720 ±0.071**	**108.73 ±23.10**	**43.11 ±3.38**	**11,822**	**149.6%**
	LR	0.514 ±0.017	143.28 ±16.57	71.02 ±3.92	20,529	363.8%
	DT	0.645 ±0.077	122.53 ±23.80	52.91 ±4.78	15,015	155.9%
	LAG_ONLY	0.425 ±0.043	155.82 ±14.54	85.25 ±5.76	24,280	694.6%
Location 2: Hämeenkatu / University corridor: device 02; main commercial street and University corridor; avg. distance = 1,088 m; n = 10,260 records.	ANN	0.832 ±0.103	121.44 ±42.47	71.53 ±22.27	14,748	44.8%
	XGB	**0.839 ±0.099**	**118.94 ±50.44**	**75.06 ±27.48**	**14,147**	**69.9%**
	LR	0.511 ±0.078	207.42 ±51.32	131.54 ±26.99	43,024	143.6%
	DT	0.537 ±0.179	201.86 ±42.35	119.05 ±24.18	40,748	66.9%
	LAG_ONLY	0.141 ±0.215	274.90 ±62.66	175.63 ±29.96	75,570	200.5%
Location 3: Logomo / Ratapihankatu: devices 13–15; cultural venue and railway district; avg. distance = 625 m; n = 25,337 records.	ANN	0.497 ±0.117	82.65 ±12.04	46.08 ±4.06	6,830	129.5%
	XGB	**0.570 ±0.105**	**76.37 ±12.29**	**42.58 ±4.00**	**5,833**	**122.0%**
	LR	0.432 ±0.041	87.83 ±13.84	56.60 ±3.76	7,714	343.2%
	DT	0.515 ±0.083	81.10 ±13.37	44.08 ±3.94	6,578	111.0%
	LAG_ONLY	0.181 ±0.057	105.46 ±13.23	72.63 ±5.88	11,122	720.4%
Location 4: Remote / Outlier: device 16; remote building entrance/exit; avg. distance = 1,529 m; n = 2,263 records.	ANN	−0.196 ±0.417	1.77 ±0.74	1.34 ±0.41	3.1	88.5%
	XGB	**0.179 ±0.231**	**1.46 ±0.67**	**1.10 ±0.30**	**2.1**	**74.5%**
	LR	−0.163 ±0.196	1.74 ±0.66	1.37 ±0.24	3.0	93.6%
	DT	−1.191 ±1.159	2.39 ±1.02	1.30 ±0.38	5.7	86.5%
	LAG_ONLY	−0.111 ±0.232	1.70 ±0.63	1.33 ±0.21	2.9	90.2%

These findings offer location-specific planning insights. High temporal predictability at the City Centre supports operational decisions such as maintenance scheduling and crowd management. The University corridor's structural patterns make it suitable for events and timetable-driven transport planning. At Logomo, the dominance of socio-economic variables points toward long-term decisions such as zoning and cycling infrastructure investment.

6 Concluding Remarks

The current research demonstrates that ML, in particular XGBoost and ANN, can accurately predict foot-traffic flows when models are tailored to the spatial and functional context of each location, as confirmed under conservative temporal evaluation. Under chronological 80/20 train–test splits with 5-fold time-series cross-validation, XGBoost achieved $R^2 = 0.720 \pm 0.071$ (City Centre) and 0.839 ± 0.099 (University corridor), confirming genuine predictive capability beyond simple autocorrelation (LAG_ONLY baseline: $R^2 = 0.425$ and 0.141, respectively; $\Delta R^2 = +0.295$ to $+0.698$). Performance at the irregular cultural venue (Location 3: $R^2 = 0.570 \pm 0.105$) and the low-volume outlier sensor (Location 4: $R^2 = 0.179 \pm 0.231$) highlights the boundary conditions of the approach. The location-stratified modelling approach, cross-modal feature relationships, and land-use sensitivity presented here provide practical guidance, scoped to the Turku context, for smart city traffic management and urban mobility planning.

Limitations include: seasonal coverage (March–November only; winter unvalidated); LSTM subsampling (full-dataset evaluation needed); single-city scope (validation in other urban contexts required); weather as observed covariate (operational deployment would require forecast values); meteorological variables (sunshine duration and cloud cover are unavailable at hourly resolution from the FMI for the study period and are not included); and evaluation uncertainty (±SD reflects fold variance, not hyperparameter uncertainty). Future work should extend seasonal coverage, incorporate event calendars and richer spatial features, and pilot the approach in a live operational setting.

Acknowledgement

The authors acknowledge the help of the City of Turku in establishing the monitoring network and thank the relevant municipal departments for providing access to the mobility dataset used in this study. The study is part of the Turku Urban Research Programme and was funded through the MUVIT project (2025-2026).

Endnotes

Data Availability: The dataset was collected in Turku, Finland, through a municipal pedestrian and traffic monitoring network. The data are available upon reasonable request from the corresponding author.

AI Disclosure: The authors used AI-assisted tools for language editing and text drafting in the preparation of this manuscript. All intellectual contributions, analyses, interpretations, and conclusions are solely those of the authors.

Conflict of Interest: The authors declare no conflict of interest.

References

Alhasnawi, B. N., Hashim, H. K., Zanker, M., & Bureš, V. (2025). The rising, applications, challenges, and future prospects of energy in smart grids and smart cities systems. Energy Conversion and Management: X, 101162.

Ali, M. (2025). Classification of imbalanced travel mode choice dataset with SMOTE and prediction using interpretable machine learning. Sustainable Futures, 10, 101119. https://doi.org/10.1016/j.sftr.2025.101119

Asher, M., Oswald, Y., & Malleson, N. (2023). Predicting pedestrian counts using machine learning. AGILE: GIScience Series, 4, 18.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Dabiri, S., & Heaslip, K. (2018). Inferring transportation modes from GPS trajectories using a convolutional neural network. Transportation Research Part C: Emerging Technologies, 86, 360–371.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv Preprint arXiv:1702.08608.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? The Journal of Machine Learning Research, 15(1), 3133–3181.

Gil, I. A., Ferreira, F. A., Ferreira, N. C., Smarandache, F., Khanam, M., & Daim, T. (2026). Artificial intelligence, society 5.0 and smart city adaptation initiatives for businesses: An integrated approach. Technovation, 150, 103377.

Haghshenas, S. S., Guido, G., Haghshenas, S. S., Astarita, V., & Simic, V. (2026). Artificial intelligence techniques in road safety for sustainable mobility: An investigation into applications, challenges, and opportunities. Reliable Decision-Making for Sustainable Transportation, 203–223.

Jamal, S., Siddiqui, F., Alam, M. A., Ayman-Mursaleen, M., Zafar, S., & Naaz, S. (2026). Deep Reinforcement Learning for Sustainable Urban Mobility: A Bibliometric and Empirical Review. Sensors (Basel, Switzerland), 26(2), 376.

Kim, S.-K., & Chan, I. C. (2025). Novel Machine Learning-Based Smart City Pedestrian Road Crossing Alerts. Smart Cities, 8(4), 114.

Koech, M. J., & Sobhani, N. (2026). Predicting Future Pedestrian Movement Patterns: An LSTM Approach Using Historical Footfall Data from Dublin. Journal of Mathematics and Data Science (JMDS), 4(1), 15–22.

Lingani, G. M., Rawat, D. B., & Garuba, M. (2019). Smart traffic management system using deep learning for smart city applications. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 0101–0106.

Murcio, R., & Wang, Y. (2025). Identify optimal pedestrian flow forecasting methods in Great Britain retail areas: A comparative study of time series forecasting on a footfall dataset. ISPRS International Journal of Geo-Information, 14(2), 50.

Ng, Y., Pei, Y., & Jin, Y. (2017). Footfall count estimation techniques using mobile data. 2017 18th IEEE International Conference on Mobile Data Management (MDM), 307–314.

Peráček, T., & Kaššaj, M. (2026). Strategic Management of Urban Services Using Artificial Intelligence in the Development of Sustainable Smart Cities—Managerial and Legal Challenges. Sustainability, 18(2), 582.

Pulugurta, S., Arun, A., & Errampalli, M. (2013). Use of artificial intelligence for mode choice analysis and comparison with the traditional multinomial logit model. Procedia-Social and Behavioral Sciences, 104, 583–592.

Schläpfer, M., Dong, L., O'Keeffe, K., Santi, P., Szell, M., Salat, H., Anklesaria, S., Vazifeh, M., Ratti, C., & West, G. B. (2021). The universal visitation law of human mobility. Nature, 593(7860), 522–527.

Wang, C., Huang, S., & Zhang, C. (2025). Short-term traffic flow prediction considering weather factors based on optimized deep learning neural networks: Bo-GRA-CNN-BiLSTM. Sustainability, 17(6), 2576.

Wolniak, R., Stecuła, K., & Aydın, B. (2024). Digital transformation of grocery in-store shopping-scanners, artificial intelligence, augmented reality and beyond: A review. Foods, 13(18), 2948.

Wu, C., Ding, H., Fu, Z., & Sun, N. (2024). Multi-Scale Spatio-Temporal Attention Networks for Network-Scale Traffic Learning and Forecasting. Sensors, 24(17), 5543. https://doi.org/10.3390/s24175543

About the authors

Katayoon Pourmahdi is a PhD researcher at the School of Business and Economics at ÅAU. Her work focuses on integrating analytical approaches into decision-making, particularly in pricing, to bridge theory and practice.

Dr **Bahareh Naseri** is a postdoctoral researcher at the School of Business and Economics at ÅAU. Her research focuses on cost–benefit analysis within value chains.

Dr **Ilia Gugenishvili** is an Assistant Professor of Marketing at the School of Business and Economics at ÅAU. His research interests include sustainability, consumer behaviour, and technology.

Dr **Jozsef Mezei** is a Professor of Information Systems at the School of Business and Economics at ÅAU. His expertise includes ML in business contexts, decision-making under uncertainty, decision-support systems, analytics, soft computing, and fuzzy logic.

Dr **Anna-Greta Nyström** is a Professor of International Business at the School of Business and Economics at ÅAU. She specializes in innovation management, business models, and market shaping in ICT and emerging technologies.

It has been proofread and complies with the language standards used in the publication.

Summary

Smart cities rely on intelligent tools to manage mobility and reduce congestion, yet context-aware prediction of pedestrian and traffic flows remains limited. This study applies five ML models to a large urban dataset from Turku, Finland, incorporating temporal, spatial, meteorological, and socio-economic factors. The results show that XGBoost and neural networks perform best, but model accuracy varies across locations. These findings highlight the need for context-aware models and offer practical insights for urban planning and real-time traffic management.