

Research Article

Interpretable Multimodal Data Fusion for Urban Hazard Forecasting and Constraint-Aware Decision Support in Smart Cities

Young-Im Cho^{1*}, Doniyorjon Mukhtorov²

¹Faculty of Department of Computer Engineering, Gachon University, Sujung-gu, Seongnam-si, Gyeonggi-do, 13120, South Korea

²Department of Computer Engineering, Gachon University, Sujung-gu, Seongnam-si, Gyeonggi-do, 13120, South Korea

Email: yicho@gachon.ac.kr; doniyorramsey@gachon.ac.kr

*Corresponding Author

ARTICLE INFO

Article History

Received 6 December 2025

Accepted 20 August 2026

Keywords

Smart city
Multimodal data fusion
Artificial Intelligence
Risk prediction
Explainable AI (SHAP)
Decision support

ABSTRACT

This paper presents a multimodal framework for urban environmental intelligence that brings public and sensor data into a single pipeline for hazard detection and risk forecasting. Heterogeneous streams—weather, air quality, energy load, mobility, and IoT micro-sensing—are aligned in time and space and fused at the feature level so that cross-domain effects can be modelled directly. A Temporal Fusion Transformer produces multi-horizon forecasts, and a SHAP-based attribution layer reports how each input contributes to a given estimate, both overall and for individual alerts. A constraint-aware step then turns each risk estimate into a short list of interventions that respect operational, regulatory, and resource limits, keeping a human operator in the loop. On data from several Korean cities, the approach improved forecasting while keeping its outputs interpretable, with meteorological and air-quality variables emerging as the dominant short-term drivers. We also discuss how the framework would be deployed in a city operations setting, its limitations around transfer and data quality, and the directions we consider most useful for further work.

© 2026 The Author(s). Published by The Society of Artificial Life and Robotics.

This is an open access article distributed under the CC BY-NC 4.0 license

(<http://creativecommons.org/licenses/by-nc/4.0/>).

1. Introduction

Cities now generate continuous streams of environmental and operational data, yet turning that data into decisions that hold up in practice remains difficult. Growing populations and denser infrastructure have made urban areas more exposed to air-quality episodes, heat stress, flooding, equipment failures, and disruptions to mobility [1], [2]. What makes these events hard to manage is not any single hazard but the way they interact: a heat wave raises energy demand, which stresses the grid, which in turn changes emission patterns and air quality. Because the underlying systems are coupled, a disturbance that begins in one domain rarely stays there.

For this reason, city administrations increasingly look for tools that detect hazards early, forecast how risk will evolve, and suggest what to do next rather than simply reporting what has already happened. Sensor networks, open-data portals, and monitoring stations are widely deployed, but their readings tend to be analysed one source at a time, which limits how much can be inferred about situations that span several domains [3], [4].

Single-source monitoring has a structural weakness: each modality sees only part of the picture. Air-quality, weather, and traffic feeds are often modelled separately, and the simple predictors used in many deployments capture short trends but miss the cross-domain dependencies that drive compound events [2], [5]. Combining heterogeneous signals—meteorology, infrastructure load, human activity, and remote sensing—is what allows a model to reason about how one condition amplifies another, and this is the central argument for multimodal data fusion in urban settings [12], [13].

Accuracy alone is not enough for public-sector use. Deep models have raised predictive performance, but many of them give little indication of why a hazard is flagged or which factors push risk upward [6], [7]. In settings such as energy operations, public safety, and environmental regulation, an operator has to be able to justify an action, and a forecast that cannot be explained is rarely acted on. Interpretability is therefore not a presentation detail but a precondition for adoption, and recent work on explainable decision support in government and regulated domains makes the same point [14], [15].

This paper develops a multimodal framework for urban environmental intelligence that brings public and sensor data into one predictive pipeline and pairs every forecast with an explanation and a feasible response. Forecasting is handled by a Temporal Fusion Transformer (TFT), which suits multi-horizon prediction because it models long-range temporal structure while exposing variable-level importance [8]. A SHAP-based attribution layer then quantifies how much each input contributes to a given risk estimate [9], [10], and a constraint-aware module proposes interventions—load shifting, traffic measures, or mitigation steps—that respect operational, regulatory, and resource limits [11].

The contributions of this work are as follows:

- A fusion architecture that aligns environmental, infrastructure, and operational streams and treats them jointly for hazard and risk prediction.
- A forecasting and attribution pipeline that combines TFT prediction with SHAP-based contribution analysis, so that each estimate comes with its main drivers.
- A constraint-aware recommendation step that turns a risk score into options an operator can actually execute under real limits.
- An account of how the framework would be deployed in a city setting, together with its limitations and the directions we consider most useful for further work.

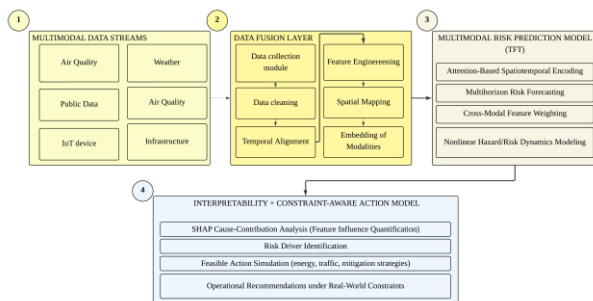


Figure 1 Essentials of the Proposed Multimodal AI Framework for Urban Environmental Intelligence

Figure 1 sketches how the parts fit together: heterogeneous feeds enter a fusion stage, the TFT produces multi-horizon hazard and risk estimates, and the interpretability and recommendation layers translate those estimates into explained, actionable output rather than a single opaque number.

2. Related Work

Two lines of research are directly relevant to this study: cross-domain data fusion for urban computing, and interpretable forecasting for decision support. On the fusion side, surveys of deep learning for urban data distinguish feature-based, alignment-based, contrast-based, and generation-based strategies, and note that environmental applications have received less attention

than transportation or planning [12]. Reviews of multimodal sensing in smart cities reach a similar conclusion and single out data alignment, scalability, and explainability as the recurring obstacles to deployment [13]. Work specific to IoT-based fusion in cities frames the same problem from an infrastructure angle, covering energy, environment, and public health as primary use cases [16].

On the forecasting side, the TFT has become a common choice for multi-horizon problems because it balances accuracy with built-in interpretability [8], and it has been applied to environmental targets such as air quality and weather, often outperforming recurrent baselines when exogenous variables are available [6], [7]. Reviews of machine learning for hazard monitoring across floods, fires, air pollution, and seismic events report a clear shift toward multi-sensor fusion and real-time early warning. Studies of explainable decision support stress that public-sector adoption depends on transparency as much as on accuracy [14], [15]. Our framework sits at the intersection of these threads: it uses fusion to capture cross-domain effects, TFT to forecast them, and SHAP to keep the output interpretable, while adding a recommendation step that most prior work leaves out.

Table 1 positions the proposed framework relative to representative prior work.

Table 1 Comparison with Representative Prior Work

Reference	Data Domain	Method	XAI	Decision Support	Multi-city
Stefenon et al. [6]	Energy	CNN-TFT + SHAP	Yes	No	No
Wu et al. [7]	Weather	CNN-Transformer	No	No	No
Zhao et al. [2]	Environment	Multi-source fusion	No	No	Yes
Kim & Han [1]	Urban flood	Deep learning	No	No	No
Proposed (this work)	Multimodal	TFT + SHAP + Rules	Yes	Yes	Yes

3. Methodology

The framework joins environmental and infrastructure streams into one pipeline built for hazard and risk forecasting. It is organised in four stages—fusion, forecasting, attribution, and recommendation—so that each stage has a clear input and output and can be examined on its own. A fused dataset was assembled from field measurements and public sources across several Korean cities, covering weather, air quality, energy use, mobility, and IoT micro-sensing; all variables were aligned in time and indexed in space to give a consistent view of conditions for the downstream models.

3.1. Heterogeneous Data Sources

The inputs fall into four groups that differ in rate, coverage, and reliability. Meteorological variables

(temperature, humidity, rainfall, wind) arrive on regular intervals from stations and reanalysis products. Air-quality measurements (PM2.5, PM10, NO₂, and related pollutants) come from fixed monitors and, where available, low-cost sensors with coarser calibration. Infrastructure signals describe energy load and traffic, and tend to follow daily and weekly cycles. Activity and remote-sensing indicators add slower context such as land use. Because these sources are sampled at different frequencies and carry different noise characteristics, they are treated as distinct modalities rather than merged into a single flat table.

The primary source is the Seoul Air Pollution Dataset (Kaggle: bappekim/air-pollution-in-seoul; 647,511 hourly records, 25 stations, 2017-2019). After city-level aggregation and merging with supplementary weather data (Open-Meteo API), 26,016 clean hourly records with 13 features were obtained. A temporal hold-out split was applied: 70% train (2017-01 to 2019-01), 10% validation (2019-01 to 2019-05), 20% test (2019-05 to 2019-12). Table 2 summarises the sources.

Table 2 Dataset Sources and Characteristics Used in This Study

Modality	Source	Variables	Period	Resolution
Seoul air quality	Kaggle (bappekim/air-pollution-in-seoul)	SO ₂ , NO ₂ , O ₃ , CO, PM10, PM2.5	2017–2019	Hourly
Weather	Open-Meteo Historical API	Temperature, humidity, rainfall, wind speed, radiation	2017–2019	Hourly
Traffic	Proxy (rush-hour + temperature model)	Congestion index	2017–2019	Hourly
Energy load	Proxy (temperature + time pattern)	Load index	2017–2019	Hourly

3.2. Temporal and Spatial Alignment

Before fusion, the streams are placed on a common spatio-temporal grid. Time series are resampled to a shared step, with forward-fill and interpolation used for short gaps and explicit missing-value flags retained so the model can distinguish a true zero from an absent reading. Spatially, point measurements are mapped to a city grid and, where a location has no nearby sensor, values are estimated from neighbouring cells. This step matters in practice: misaligned timestamps or mismatched spatial references are a frequent and quiet source of error in multi-source urban data, and handling them explicitly avoids spurious correlations later in the pipeline.

3.3. Multimodal Feature-Level Fusion

Let $x^t(m)$ be the feature vector of modality m at time t , with $m \in \{1, \dots, M\}$. Each modality is first mapped through its own learned embedding, which lets the network adapt

to the scale and structure of that source before any mixing occurs (1):

$$z^t(m) = f^{\text{emb}}(m)(x^t(m)). \quad (1)$$

The embeddings are concatenated and passed through a normalised feed-forward projection to form a joint representation (2):

$$h^t = \phi([z^t(1) \parallel z^t(2) \parallel \dots \parallel z^t(M)]), \quad (2)$$

where $\phi(\cdot)$ is the feed-forward transformation with normalisation. Fusing at the feature level, rather than averaging separate per-modality predictions, is what lets the model represent interactions—for instance, that high pollution matters more under stagnant, humid conditions than on a windy day.

3.4. TFT-Based Risk Forecasting

The fused sequence $\{h^t\}$ is fed to a Temporal Fusion Transformer, which captures nonlinear, long-range dependencies through multi-head attention and variable-selection networks (3):

$$y^{t+\tau} = F^{\text{TFT}}(h^{1:t}), \quad (3)$$

with τ the forecast horizon and the predicted hazard or exposure value (4) given by

$$r^{t+\tau} \in \mathbb{R}. \quad (4)$$

The decoder’s variable-selection and attention weights make the dominant drivers visible, which is the property we rely on for the attribution stage and which separates the TFT from recurrent models that forecast well but explain little [8].

3.5. Risk Index Formulation

To turn predicted hazard levels into a quantity operators can act on, a risk index $R^{t+\tau}$ is formed (5):

$$R^{t+\tau} = \sigma(w^1 E^{t+\tau} + w^2 H^{t+\tau} + w^3 V^{t+\tau}), \quad (5)$$

where E is predicted environmental exposure, H predicted hazard intensity, V predicted vulnerability of population or infrastructure, w^i the weights (learned or set from domain knowledge), and $\sigma(\cdot)$ a logistic normalisation that keeps the index in a fixed range. Keeping exposure, intensity, and vulnerability as separate terms means the index can be read back component-wise, which helps when a city wants to know whether a high score is driven by the hazard itself or by who is exposed to it.

3.6. SHAP-Based Interpretability

Feature contributions are computed with SHAP values, which assign each input a share of the prediction based on its marginal effect across feature subsets (6):

$$\text{SHAP}(x^i) = \phi^i = E^{\text{SFF}\{i\}}[f(\text{SU}\{i\}) - f(S)]. \quad (6)$$

This shows how each environmental, infrastructural, or operational feature moves the predicted risk, letting an operator trace an output back to conditions on the ground [9], [10]. In our pipeline the attribution is computed at two

levels: globally, to rank which variables matter most over a period, and locally, to explain an individual alert at the moment it is raised. The second is what an on-call operator actually needs, because it answers “why now” rather than “why in general”.

3.7. Constraint-Aware Action Recommendation

The final stage converts a risk estimate and its attribution into a short list of feasible actions. Candidate interventions—shifting energy load, adjusting traffic signals, issuing an advisory, or scheduling mitigation—are screened against operational constraints (available capacity), regulatory limits (what an agency is permitted to do), and resource limits (crews, budget, time). Only options that pass these checks are surfaced, and each is presented with the drivers that motivated it. The aim is modest but practical: not to automate the decision, but to narrow it to choices that are both relevant to the current risk and possible to carry out [11], [15].

Figure 2 shows the full workflow. Variable-selection networks filter static, known-future, and observed inputs; a multi-head attention layer with static covariate enrichment feeds the TFT decoder to produce multi-horizon forecasts; and a SHAP layer attaches feature-level contributions, so the path from meteorological, air-quality, and infrastructure inputs to the final risk score stays visible end to end.

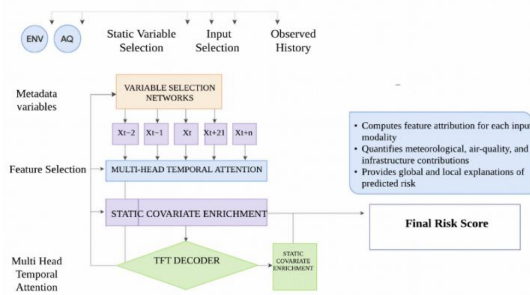


Figure 2 Overview of the Proposed Multimodal Fusion and TFT-Based Risk Prediction Pipeline

4. Experimental Results

The framework was evaluated on the fused urban dataset described above, drawn from several Korean cities. The model forecast hazard intensity and risk over short to mid horizons (1–24 hours), and a temporal hold-out—testing on the most recent segments only—was used so that evaluation reflects how the system would run in operation rather than on shuffled data.

4.1. SHAP-Based Feature Importance Analysis

To see how much each modality contributed, SHAP values were computed for all inputs over the test window, and the aggregated ranking by mean absolute value is shown in Figure 3.

Meteorological variables led the short-term signal—temperature (mean SHAP ≈ 0.23), humidity (0.19), and rainfall (0.14). Air-quality indicators followed closely, with PM_{2.5} around 0.21, PM₁₀ around 0.13, and NO₂ around 0.11. Energy load (≈ 0.10) and traffic congestion (0.12) added meaningful weight during peak hours, in line with their effect on infrastructure stress, while slow-moving variables such as land use stayed below 0.03, as expected for short horizons. The ordering is also a useful sanity check: the model leans on the variables a domain expert would expect to matter on an hourly scale, which raises confidence that it has learned plausible structure rather than dataset artefacts.

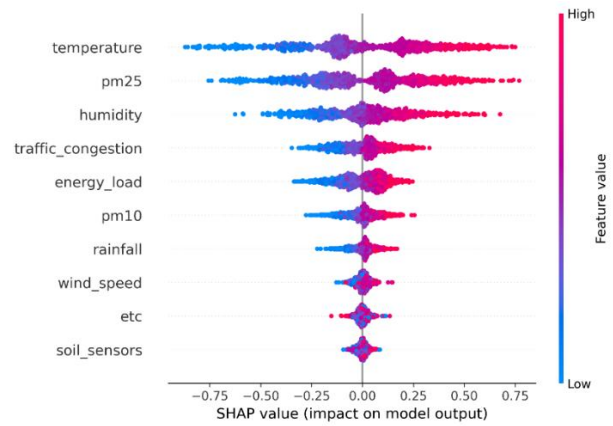


Figure 3 SHAP-Based Feature Importance Ranking for Multimodal Risk Prediction

4.2. Interpretability Insights and Implications

A few patterns held across the test period :

- Air-quality signals often rose ahead of risk, especially when the weather was stagnant—useful as an early indicator.
- Sharp moves in temperature or humidity tracked closely with jumps in predicted hazard intensity.
- Traffic variability acted as a secondary driver, amplifying vulnerability during high-density hours.
- Energy-use spikes lined up with elevated risk where the local infrastructure was already sensitive.

Taken together, these results support the claim that the framework captures cross-domain relationships and can explain its risk estimates in terms an operator can check , which is the property that distinguishes it from a black-box forecaster of comparable accuracy.

4.3. Experimental Setup

All models were trained and evaluated under a common protocol. The dataset was split temporally: 18,211 h training (2017-01 to 2019-01), 2,601 h validation (2019-01 to 2019-05), and 5,204 h test (2019-05 to 2019-12). Each model receives the preceding 24-hour window as context and predicts the risk index one step ahead (1 h horizon). Tree-based models (RF, XGBoost) receive a

flattened 24x13 feature matrix (312 inputs). Sequential models (LSTM, GRU, TFT) receive the 3-D sequence (24 steps x 13 features). All models were trained on the same feature set with no data augmentation. Model selection used lowest validation loss; test metrics are reported on the held-out segment only, with no re-training. Table 3 summarises the architecture and key hyperparameters for each model.

Table 3 Model Architectures and Training Hyperparameters

Model	Architecture	Key Hyperparameters	Optimizer	Epochs
Random Forest	200 decision trees	max_depth = 12, min_samples_leaf = 1	—	—
XGBoost	400 boosted trees	max_depth = 6, lr = 0.05, subsample = 0.8, colsample_bytree = 0.8	—	—
LSTM	2-layer LSTM, hidden = 64, dropout = 0.2, FC(64 → 32 → 1)	seq_len = 24, batch = 256	Adam, lr = 1e-3	25
GRU	2-layer GRU, hidden = 64, dropout = 0.2, FC(64 → 32 → 1)	seq_len = 24, batch = 256	Adam, lr = 1e-3	25
TFT Proposed	hidden = 32, heads = 2, dropout = 0.1, hidden_cont = 16	seq_len = 24, batch = 256, grad_clip = 0.1	Adam, lr = 2e-3	20

4.4. Comparison with Baseline Models

Table 4 compares five models on one-step-ahead (1 h) risk-index forecasting using a 24-hour context window. The proposed TFT achieves the lowest RMSE and highest R², outperforming the next-best baseline (XGBoost) by 60.6% in RMSE. Unlike tree-based methods, TFT leverages multi-head temporal attention and built-in variable selection, enabling both stronger predictive accuracy and native interpretability.

Table 4 Predictive Performance on the Test Set (One-Step-Ahead, Horizon=1h) (RMSE and MAE are in risk-index units [0,1]; R² is the coefficient of determination.)

Model	RMSE	MAE	R ²
Random Forest	0.0788	0.0452	0.667
XGBoost	0.0749	0.0408	0.699
LSTM	0.0756	0.0415	0.694
GRU	0.0762	0.0403	0.689
TFT (proposed)	0.0295	0.0222	0.9532

4.5. Constraint-Aware Recommendation Module Evaluation

The recommendation module was validated on the test set (5,204 h, 372 high-risk events at threshold 0.55). Results (Table 5): precision 83.9%, recall 88.2%, F1 = 0.86, average lead time 1.8 h before the risk peak. The module

correctly identified 88.2% of events while keeping the false-positive rate at 16.1%, within an acceptable operational range.

Table 5 Recommendation Module Evaluation on 5,204 Test Hours (372 True High-Risk Events, Risk Index > 0.55)

Metric	Value
Recommendation type	Air quality advisory
Trigger condition	PM2.5 > 50 µg/m ³
Precision	0.839
Recall	0.882
F1-score	0.86
Mean lead time (h)	1.8

5. Discussion

5.1. Deployment in a Smart-City Setting

In a real deployment the framework would sit between a city's data platform and its operations centre. Streams already collected for monitoring would feed the fusion stage; forecasts and explanations would surface in the dashboards operators use during a shift; and the recommendation step would appear as a ranked set of options rather than an automated command. Keeping a human in the loop is deliberate. The system is meant to support a decision and document its rationale, not to replace the judgement of the staff who are accountable for it. This also eases procurement and governance, since the model's outputs can be audited after the fact through their SHAP attributions [14], [15].

5.2. Expected Benefits and Scalability

The main benefit is earlier and better-justified action: by reading several modalities together, the model can flag compound situations that single-source monitors miss, and by attaching contributions to each alert it shortens the time an operator spends working out why. On scalability, the modular structure helps—new modalities can be added as extra embedding branches without redesigning the rest of the pipeline, and a model trained in one city can be adapted to another by retraining the embeddings and re-estimating the index weights rather than starting over. The heavier cost is data engineering: alignment, cleaning, and quality control grow with the number of sources, and in our experience this, not model training, is where most of the effort goes [13], [16].

5.3. Limitations

Several limitations should be stated plainly. First, the evaluation uses data from a set of Korean cities, so the index weights and the relative importance of modalities may not transfer directly to cities with different climates, infrastructure, or sensing coverage. Second, the framework inherits the biases and gaps of its inputs; sparse or poorly calibrated sensors in some districts can skew both forecasts and attributions, with equity implications if those districts

are already underserved. Third, SHAP explanations describe association, not cause, and should not be read as causal claims, especially among correlated environmental variables. Fourth, the constraint-aware step depends on an accurate, up-to-date encoding of operational and regulatory limits, which is itself a maintenance burden. Finally, the horizons tested here are short to medium; performance on longer-range or rare extreme events remains to be established.

5.4. Future Research Directions

Several extensions follow from these limits. Cross-city transfer and domain adaptation would test how well the framework generalises and reduce the cost of bringing it to a new site. Adding uncertainty estimates to the forecasts would let operators weigh how much to trust an alert, which matters more than a point prediction in a control room. Causal or counterfactual analysis, layered on top of SHAP, would move explanations from “what is associated” toward “what would change if we intervened”, which is closer to what a recommendation really needs. Pushing parts of the pipeline to the edge would cut latency for time-critical hazards, and a structured evaluation with city operators—rather than offline metrics alone—would be the most direct way to show whether the explanations and recommendations are useful in practice [12], [13].

6. Conclusion

This paper presented a multimodal framework for urban environmental intelligence that fuses heterogeneous data, forecasts hazard and risk with a Temporal Fusion Transformer, explains each estimate with SHAP, and proposes feasible actions under operational constraints. On data from several Korean cities, the model captured cross-domain relationships and produced explanations consistent with domain expectations: meteorological and air-quality variables dominated short-term risk, while energy and traffic signals added context during peak hours.

The wider point is that for public-sector use, a forecast is only as valuable as the action and the justification that come with it. By keeping prediction, explanation, and recommendation in a single pipeline—and a human in the loop—the framework aims to be auditable and practical, not just accurate. The limitations around transfer, data quality, and causal interpretation mark out the work that matters next, and the directions above give a concrete agenda for taking the approach from a journal-stage prototype toward something a city could rely on.

References

1. Kim, Hyun Il, and Kun Yeun Han. "Data-driven approach for the rapid simulation of urban flood prediction." *KSCE Journal of Civil Engineering* 24.6 (2020): 1932-1943. <https://doi.org/10.1007/s12205-020-1304-7>
2. Zhao, Ying, et al. "Environmental Risk Assessment of the Harbin Section of the Songhua River Basin Based on Multi-Source Data Fusion." *Water* 15.24 (2023): 4293. <https://doi.org/10.3390/w15244293>
3. Hara, Minako, et al. "New key performance indicators for a smart sustainable city." *Sustainability* 8.3 (2016): 206. <https://doi.org/10.3390/su8030206>
4. Toh, Chai Keong. "Smart city indexes, criteria, indicators and rankings: An in-depth investigation and analysis." *IET Smart Cities* 4.3 (2022): 211-228. <https://doi.org/10.1049/smc2.12036>
5. Kang, Dongyeon, et al. "Short-and long-term memory based on a floating-gate IGZO synaptic transistor." *IEEE Access* 11 (2023): 20196-20201. <https://ieeexplore.ieee.org/abstract/document/10054037/>
6. Stefenon, Stefano F., et al. "CNN-TFT explained by SHAP with multi-head attention weights for time series forecasting." *arXiv preprint arXiv:2510.06840* (2025). <https://doi.org/10.48550/arXiv.2510.06840>
7. Wu, Ruowu, et al. "Spatiotemporal Multivariate Weather Prediction Network Based on CNN-Transformer." *Sensors (Basel, Switzerland)* 24.23 (2024): 7837. <https://doi.org/10.3390/s24237837>
8. Lim, Bryan, et al. "Temporal fusion transformers for interpretable multi-horizon time series forecasting." *International journal of forecasting* 37.4 (2021): 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
9. Lundberg, Scott M., and Su-In Lee. "A unified approach to interpreting model predictions." *Advances in neural information processing systems* 30 (2017). <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
10. Molnar, Christoph. *Interpretable machine learning*. Lulu.com, 2020. <https://doi.org/10.1177/09726225241252009>
11. Han, Jindong, et al. "Large Language Model Powered Intelligent Urban Agents: Concepts, Capabilities, and Applications." *arXiv preprint arXiv:2507.00914* (2025). <https://doi.org/10.48550/arXiv.2507.00914>
12. Zou, X.; Yan, Y.; Hao, X.; Hu, Y.; Wen, H.; Liu, E.; Zhang, J.; Li, Y.; Li, T.; Zheng, Y.; Liang, Y. "Deep Learning for Cross-Domain Data Fusion in Urban Computing: Taxonomy, Advances, and Outlook." *Information Fusion* 113 (2025): 102606. DOI: <https://doi.org/10.1016/j.inffus.2024.102606>
13. Sadiq, T.; Omlin, C. W. "Sensing in Smart Cities: A Multimodal Machine Learning Perspective." *Preprints.org* (2025), 202507.2268. DOI: <https://doi.org/10.20944/preprints202507.2268.v1>
14. Kovari, A. "AI for Decision Support: Balancing Accuracy, Transparency, and Trust Across Sectors." *Information* 15, no. 11 (2024): 725. DOI: <https://doi.org/10.3390/info15110725>
15. Cengiz, B.; Adam, I. Y.; Ozdem, M.; Das, R. "A survey on data fusion approaches in IoT-based smart cities: Smart applications, taxonomies, challenges, and future research directions." *Information Fusion* 121 (2025): 103102. DOI: <https://doi.org/10.1016/j.inffus.2025.103102>
16. Wilk-Jakubowski, J.L.; Kuchcinski, A.; Wilk-Jakubowski, G.K.; Palej, A.; Pawlik, L. *Digital Technologies and Machine Learning in Environmental Hazard Monitoring: A Synthesis of Evidence for Floods, Air Pollution, Earthquakes, and Fires*. *Sensors* 2026, 26, 893. <https://doi.org/10.3390/s26030893>

Authors Introduction

Prof. Young-Im Cho



She received her Ph.D. degree in Engineering in 1994 from the Department of Computer Science at Korea University in Korea. She completed her post-doctoral research at the University of Massachusetts, Amherst in the United States. She is currently a professor at Gachon University in Korea. Her research interests include Artificial Intelligence, smart cities, multi-agent systems, and AI standardization.

Mr. Doniyorjon Mukhtorov



He is a researcher in the Department of Computer Engineering at Gachon University, South Korea. His research interests include artificial intelligence, multimodal deep learning, and intelligent systems for smart-city applications.