A Glass-Box AI Framework for Predicting and Explaining Urban Solar Energy Generation

Yumin Zhou
College of Computing & Data Science
Nanyang Technological University
50 Nanyang Ave, Singapore 639798
s230038@e.ntu.edu.sg

Whye Loon Tung
School of Business
Singapore University of Social Sciences
463 Clementi Rd, Singapore 599494
wltung001@suss.edu.sg

Quek Hiok Chai
College of Computing & Data Science
Nanyang Technological University
50 Nanyang Ave, Singapore 639798
ashcquek@ntu.edu.sg

Abstract

The reliable integration of volatile solar power is critical for sustainable urban energy systems. Engineers face a choice between transparent classical models that lack accuracy and high-performing black-box' models that pose operational risks. This paper breaks that trade-off by introducing L-FMLC, a 'glass-box' AI framework that delivers both state-of-the-art performance and deep interpretability. L-FMLC provides instance-specific linear equations for every prediction, grounded in autonomously discovered fuzzy sets that represent intuitive concepts like 'clear' and 'overcast' skies. It then distills thousands of potential decision paths into a handful of strategic 'meta-rules', culminating in natural-language reports. In a real-world solar forecasting case study, we show L-FMLC surpasses classical and standard deep learning baselines while providing this full stack of actionable explanations. This work offers a practical blueprint for building AI systems for urban infrastructure that are simultaneously high-performing and fundamentally trustworthy.

1 Introduction

Integrating volatile solar power into urban energy grids is a critical challenge for sustainable cities. Urban planners and engineers have historically faced a difficult trade-off when selecting forecasting tools. They could choose classical empirical models, which offer transparent equations but are often brittle, relying on few inputs and failing to capture local dynamics [1]. Or, they could opt for modern deep learning models like LSTMs [2] and Transformers [3], which achieve impressive accuracy by learning from dozens of variables but operate as inscrutable 'black boxes'. For critical infrastructure, deploying a model that an operator cannot trust, query, or understand is an unacceptable risk [4, 5].

This paper argues that this trade-off is no longer necessary. We introduce the **Learnable Fuzzy-Modulated Linear Consequents (L-FMLC)** framework, a 'glass-box' AI system we developed to synergize the predictive power of deep learning with the structured logic of fuzzy systems [6]. This synergy results in a framework that achieves both state-of-the-art accuracy and deep, multi-level interpretability.

In this work, we demonstrate how L-FMLC provides a complete solution through three key contributions:

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: UrbanAI: Harnessing Artificial Intelligence for Smart Cities.

(i) A transparent architecture that surpasses both classical and black-box baselines in predictive accuracy.

(ii) A mechanism for autonomously discovering tangible system states, such as distinct weather conditions, which form the basis for instance-level explanations.

(iii) A complete pipeline to manage, distill, and translate thousands of low-level TSK rules into a compact set of strategic meta-rules and natural-language reports.

2 The L-FMLC Glass-Box Framework

L-FMLC [6] is a neuro-fuzzy system architected for both high-performance prediction and deep interpretability. As illustrated in Figure 1, it operates via two integrated pathways: a top-down process for generating predictions and a bottom-up pipeline for extracting human-understandable rules.

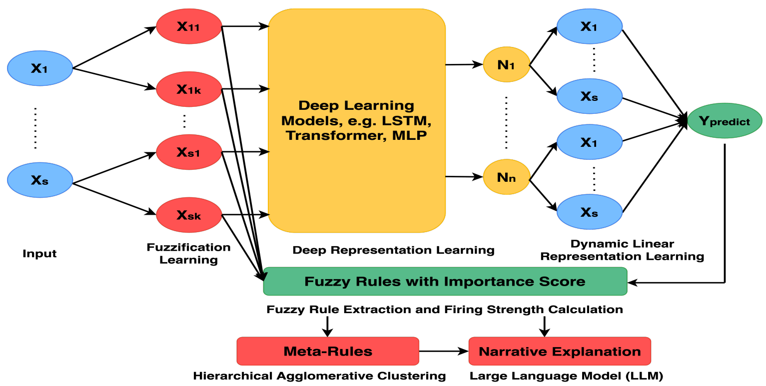

Figure 1: The dual-path architecture of L-FMLC. The prediction pathway (top) generates accurate forecasts using a deep model to dynamically parameterize a linear layer. The parallel interpretation pipeline (bottom) distills the model's complex logic into a handful of meta-rules, culminating in natural language explanations from an Large Language Model (LLM).

Adaptive Fuzzification. The framework's entry point is a fully adaptive fuzzification layer that transforms raw inputs (X) into a meaningful, problem-specific representation. Unlike static approaches that use fixed partitions [7], L-FMLC learns the optimal parameters of its Gaussian membership functions (centroids c, widths σ) end-to-end. More importantly, it autonomously determines the necessary *number* of fuzzy sets for each feature. This is achieved by training learnable gates, $g_{s,k}$, for each potential set, guided by a composite regularization scheme that encourages sparsity (pruning unneeded sets), distinctness (preventing semantic overlap), and data coverage [6]. This process yields a parsimonious and interpretable set of learned concepts tailored to the task at hand.

Dynamic Prediction Pathway. Once the inputs are fuzzified, a *Deep Representation Learner* (e.g., a Transformer) processes this rich representation to produce a small vector of context 'modulators' (N). These modulators are the key to the model's dynamic reasoning. They are fed to the final linear layer, where they construct instance-specific coefficients (C_s) for each of the original input features:

$$y_p = \sum_{s=1}^{S} (C_s(N) \times X_s)$$

This structure, first proposed in FMLC [7], is inherently explanatory. For any prediction, one can inspect the exact linear equation the model used, revealing precisely how each input feature was weighted in that specific context.

Scalable Rule Management. The interpretability pipeline is designed to manage the combinatorial explosion of potential fuzzy rules. First, we can extract TSK-style fuzzy rules that describe the model's average linear logic within specific fuzzy regions [7]. However, this can yield hundreds of rules, making direct interpretation impractical. To solve this, our L-FMLC framework [6] employs a two-stage distillation process. First, the adaptive fuzzification layer provides *inherent architectural pruning* by learning a minimal set of necessary fuzzy concepts. Second, we use Hierarchical

Agglomerative Clustering [8] on the remaining TSK rules, grouping them into a handful of 'meta-rules' that represent the model's core operational logic. This process culminates in a final translation step, where an LLM converts the structured meta-rules into natural-language narratives, making the model's core logic accessible to any stakeholder.

3 Case Study: Solar Irradiation Prediction

We validate L-FMLC's utility in a real-world urban scenario: forecasting solar generation using data from the National Solar Radiation Database (NSRDB) [9]. The objective is to predict Global Horizontal Irradiance (GHI) from 20 meteorological features. We rigorously evaluate our framework against a spectrum of alternatives, from classical equations to modern black-box models.

Baselines for Comparison. We benchmark L-FMLC against three key models. (1) A **Classical Empirical Baseline**, for which we implemented the final, locally-tuned equation from Yang et al. [1]. This allows a direct, fair comparison against traditional physics-informed regression on our own dataset. (2) A **Standard Deep Model**, a standard Transformer architecture [10] trained on the raw features, representing a powerful but opaque baseline. (3) The **Static FMLC** framework [7], which uses a non-learnable fuzzification layer, to isolate the benefits of our adaptive approach.

Performance for Critical Infrastructure. For any model to be trusted, it must first be accurate. Table 1 reveals a clear performance hierarchy. Our full, regularized L-FMLC framework achieves the best results across all metrics, demonstrating its guided, adaptive structure finds a more robust and generalizable solution.

Table 1: Predictive Performance on Solar Irradiation. Our L-FMLC outperforms both classical empirical and modern deep learning baselines.

Model	**R^2 (↑)**	**RMSE (↓)**	**MAPE(%) (↓)**
L-FMLC-Transformer (w/ Reg)	**0.8852**	**91.33**	**0.0589**
Static FMLC-Transformer [7]	0.8621	96.27	0.0642
Transformer (Standard Deep Model) [10]	0.6880	142.32	0.2161
Yang et al. (2012) [Empirical] [1]	0.5731	215.82	0.3540

Actionable Insights from the Glass Box. Beyond its accuracy, L-FMLC's value lies in its multi-level interpretability, which provides different insights for different stakeholders. **Level 1** offers instance-specific equations for the system operator validating a live forecast. **Level 2** reveals the model's standard logic for recurring scenarios (e.g., 'all overcast days'), which is essential for a diagnostics engineer. **Level 3** provides strategic 'meta-rules' for the system planner analyzing long-term behavior. Finally, **Level 4** generates natural-language reports for non-technical managers.

Level 1: Instance Explanations via Learned Concepts. The most granular insight comes from the model's dynamic output layer, which generates a precise linear equation for every prediction, such as: $y_p = 0.89 \times (\text{CSGHI}) - 0.21 \times (\text{Cloud}) \ldots$. This is enabled by an adaptive front-end that autonomously discovers fuzzy concepts from data. Figure 2 shows a clear example for the 'Total Cloud Cover' feature. Instead of relying on a pre-defined partition (a), L-FMLC autonomously learns that an optimal structure requires only three fuzzy sets (b). These learned sets align perfectly with the intuitive domain concepts of 'Clear', 'Partly Cloudy', and 'Overcast' skies—all discovered from data without human supervision. It is these learned concepts that determine the instance-specific coefficients in the final equation, allowing an operator to validate the model's reasoning on a case-by-case basis.

Level 2: Region-Specific TSK Rules. While Level 1 explains individual predictions, we can generalize to understand the model's average behavior across fuzzy regions (**Level 2**). This involves extracting TSK-style fuzzy rules that define the model's typical linear logic for a given scenario [7], such as for all instances where 'Cloud Cover' is 'High' and 'Zenith Angle' is 'High':

IF Cloud Cover is *High* AND Zenith Angle is *High* ...
THEN $y_p = 0.15 \times (\text{Clear Sky GHI}) - 0.55 \times (\text{Cloud Cover}) \ldots$

Level 3: High-Level Strategic Meta-Rules. However, Level 2 still yields 298 distinct TSK rules for our case study—far too many for direct human interpretation. Therefore, we use hierarchical

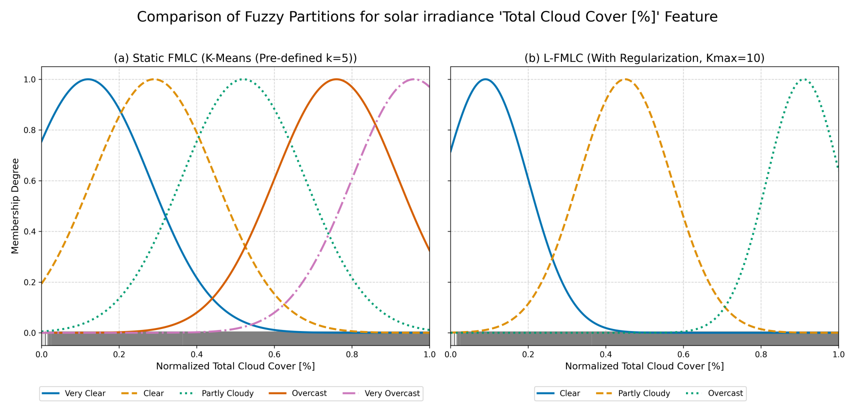

Figure 2: **Autonomous fuzzy set discovery for 'Total Cloud Cover'.** (a) A static, pre-defined partition is suboptimal. (b) L-FMLC automatically learns a parsimonious K=3 structure, with the learned sets corresponding directly to the intuitive concepts of 'Clear', 'Partly Cloudy', and 'Overcast' skies.

clustering to distill the entire rule base into the five core 'meta-rules' shown in Table 2. This step transforms a complex set of behaviors into a concise diagnostic map of the system's primary operational modes.

Table 2: Distilled Meta-Rules for the L-FMLC Solar Model.

Meta-Rule	**Qualitative Description (Dominant Antecedent Logic)**	**Raw Rules**
1. Low Light	IF Solar Zenith Angle is *High* AND Cloud Cover is *High/Overcast* THEN Output is **Low**	45
2. Ideal Sunny Day	IF Solar Zenith Angle is *Low* AND Clear Sky Index is *High* AND Cloud Cover is *Low* THEN Output is **High**	90
3. Hot & Overcast	IF Temp is *High* AND Humidity is *High* AND Cloud Cover is *Mostly Overcast* THEN Output is **Med-Low**	55
4. Cold & Clear	IF Solar Zenith Angle is *High* AND Clear Sky Index is *High* AND Temp is *Low* THEN Output is **Medium**	63
5. Scattered Clouds	IF Cloud Cover is *Partly Cloudy* AND GHI is *Medium* THEN Output is **Med-High**	45

Level 4: Natural Language Reports. Finally, the pipeline culminates in a translation step where the structured logic of a meta-rule is fed as a prompt to an LLM. This use of language models as explainers [11] automatically converts the quantitative reasoning into a fluid, plain-language summary (Figure 3), making the model's core logic accessible to any stakeholder.

Input to LLM (Prompt):

```
# Explain this TSK fuzzy rule from a solar prediction model
# Context: Predicts solar irradiation (GHI).
# Meta-Rule: "Ideal Sunny Day"
# ANTECEDENT (Dominant Conditions):
# - Solar Zenith Angle is 'Low'
# - Clear Sky Index is 'High'
# - Cloud Cover is 'Low'
# CONSEQUENT (Key Averaged Coefficients):
# - Coeff for Clear Sky GHI: +0.85
# - Coeff for Cloud Cover: -0.15
# - Coeff for Solar Zenith Angle: -0.45
```

Output from LLM (Gemini-2.5-Pro):

This rule describes when the sun is high in the sky (low zenith angle) and the air is very clear of haze, and there are almost no clouds, the model predicts a high level of solar irradiation. The most important factor in this scenario is the amount of direct sunlight available, which has a very strong positive influence (+0.85) on the prediction. The sun's exact position (zenith angle), which always affects radiation intensity. A low (negative) solar zenith angle corresponds to a high position of the sun in the sky, the negative coefficient of -0.45 creates a significant positive contribution to the final prediction. Interestingly, The model also notes that while cloud cover is low, it still has a minor negative impact.

Figure 3: An L-FMLC meta-rule is translated into a fluid natural language explanation by an LLM, making quantitative logic accessible.

4 Discussion and Conclusion

Our solar forecasting case study proves that the trade-off between performance and transparency in urban AI is no longer a given. The L-FMLC framework delivers state-of-the-art accuracy while providing full stack of actionable insights—a principle that extends directly to other critical domains like building energy optimization and traffic management, where a model's ability to explain its reasoning is paramount for operator trust and fault diagnosis.

The critical next step for real-world deployment is robust uncertainty quantification. For an urban grid operator, knowing a model's confidence ('I am 95% sure the output will be between X and Y') is as crucial as the prediction itself, enabling true risk-aware decision-making, such as deploying backup power only when a solar forecast is uncertain, further cementing trust in high-stakes operations.

Ultimately, the L-FMLC 'glass-box' approach offers a practical blueprint for developing the trustworthy, high-performance AI needed to build the resilient and sustainable urban systems of the future.

References

[1] Y. Dazhi, P. Jirutitijaroen, and W. M. Walsh, "The estimation of clear sky global horizontal irradiance at the equator," *Energy Procedia*, vol. 25, pp. 141–148, 2012, pV Asia Pacific Conference 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1876610212011812

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 11 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2023. [Online]. Available: https://arxiv.org/abs/1706.03762

[4] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information Fusion*, vol. 58, pp. 82–115, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253519308103

[5] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM Comput. Surv.*, vol. 51, no. 5, Aug. 2018. [Online]. Available: https://doi.org/10.1145/3236009

[6] Anonymous, "L-FMLC: End-to-end neuro-fuzzy learning for adaptive and scalable interpretability," 2025, under review.

[7] ——, "Fuzzy-modulated linear consequents for enhanced performance and interpretability in large models," 2025, under review.

[8] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: An overview," *Wiley Interdisc. Rew.: Data Mining and Knowledge Discovery*, vol. 2, pp. 86–97, 01 2012.

[9] M. Sengupta, Y. Xie, A. Habte, A. Lopez, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 51–60, 2018.

[10] B. Xiong, Y. Chen, D. Chen, J. Fu, and D. Zhang, "Deep probabilistic solar power forecasting with transformer and gaussian process approximation," *Applied Energy*, vol. 382, p. 125294, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261925000248

[11] D. Martens, J. Hinns, C. Dams, M. Vergouwen, and T. Evgeniou, "Tell me a story! narrative-driven xai with large language models," *Decision Support Systems*, vol. 191, p. 114402, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016792362500003X

[12] G. V. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, pp. 303–314, 1989. [Online]. Available: https://api.semanticscholar.org/CorpusID:3958369

[13] L. Bottou, F. E. Curtis, and J. Nocedal, “Optimization methods for large-scale machine learning,” *SIAM Review*, vol. 60, no. 2, pp. 223–311, 2018. [Online]. Available: https://doi.org/10.1137/16M1080173

[14] S. J. Reddi, S. Kale, and S. Kumar, “On the convergence of adam and beyond,” 2019. [Online]. Available: https://arxiv.org/abs/1904.09237

[15] Y. Finance. (2025) Yahoo finance sp 500 stock price hisotry. [Online]. Available: https://finance.yahoo.com/quote/%5EGSPC/history/

[16] M. Zwitter and M. Soklic, “Breast Cancer,” UCI Machine Learning Repository, 1988, DOI: https://doi.org/10.24432/C51P4M.

[17] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2017. [Online]. Available: https://arxiv.org/abs/1412.6980

A Technical Appendices and Supplementary Material

This appendix provides supplementary details for our neuro-fuzzy framework, including detailed mathematical formulations, experimental setup, and additional results.

A.1 Framework Details

A.1.1 Learnable Fuzzification and Regularization Loss

As described in the main paper, the learnable fuzzification layer uses gated Gaussian membership functions (MFs). The final gated membership degree $\hat{\mu}_{s,k}$ for feature s and concept k is:

$$\hat{\mu}_{s,k}(x_s) = g_{s,k} \cdot \exp\left(-\frac{(x_s - c_{s,k})^2}{2\sigma_{s,k}^2}\right) \tag{1}$$

where the center $c_{s,k}$, standard deviation $\sigma_{s,k}$, and gate $g_{s,k}$ are all learnable parameters.

The total loss function used to train the model is a combination of the primary task loss ($\mathcal{L}_{\text{task}}$) and the three cognitive regularizers:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{task}} + \lambda_{\text{sparse}}\mathcal{L}_{\text{sparse}} + \lambda_{\text{overlap}}\mathcal{L}_{\text{overlap}} + \lambda_{\text{coverage}}\mathcal{L}_{\text{coverage}} \tag{2}$$

where λ are hyperparameter weights. The specific formulations for the regularization terms are:

- **Parsimony ($\mathcal{L}_{sparse}$):** An L1 penalty on the gates: $\mathcal{L}_{\text{sparse}} = \sum_{s,k} |g_{s,k}|$.
- **Distinctness ($\mathcal{L}_{overlap}$):** A penalty based on the intersection of adjacent active MFs: $\mathcal{L}_{\text{overlap}} = \sum_{s,k \in \text{active}} \exp\left(-\frac{(c_{s,k+1} - c_{s,k})^2}{\sigma_{s,k}^2 + \sigma_{s,k+1}^2}\right)$.
- **Coverage ($\mathcal{L}_{coverage}$):** A penalty for MFs not spanning the data range: $\mathcal{L}_{\text{coverage}} = \sum_s ((\min_k c_{s,k} - x_{s,\min})^2 + (\max_k c_{s,k} - x_{s,\max})^2)$.

A.1.2 Dynamic Consequent Mechanism

The core reasoning process (Eq. 2 in the main paper) is instantiated by a deep model (`DeepModel`) that takes the fuzzified vector M as input to produce n_N modulators $N = [N_1, \ldots, N_{n_N}]^T$. These are combined with a trainable weight matrix $W \in \mathbb{R}^{S \times n_N}$ to compute the dynamic coefficients C_s:

$$C_s(\text{context}) = \sum_{i=1}^{n_N} W_{s,i} \cdot N_i \tag{3}$$

This allows the final prediction to be a transparent, instance-specific linear function of the original inputs.

A.1.3 Rule Vectorization for Clustering

To perform hierarchical clustering on the extracted TSK rules, each rule is converted into a single numerical vector. For a rule defined by an antecedent partition P and its characteristic consequent, its vector representation V_P is formed by concatenating:

1. **The Antecedent Vector:** The learned parameters of the MFs defining the antecedent: $[c_{1,k_1}, \sigma_{1,k_1}, \ldots, c_{S,k_S}, \sigma_{S,k_S}]$.
2. **The Consequent Vector:** The vector of characteristic TSK coefficients: $[C_{P,1}, \ldots, C_{P,S}]$.

The entire dataset of rule vectors is then standardized before applying Hierarchical Agglomerative Clustering using Ward's linkage.

A.2 Universal Approximation Capacity of L-FMLC

This section provides a more detailed proof for the universal approximation capability of the L-FMLC framework, which was stated in Theorem 1.

Theorem 1 (Universal Approximation for L-FMLC)**.** *Let an L-FMLC model be defined as in Section 3 The L-FMLC Framework. Assume its membership functions are continuous and differentiable with respect to their learnable parameters, and its* `DeepModel` *is a universal approximator. For any target function $f(X)$ on a compact domain $D \subset \mathbb{R}^S$ that can be expressed as $f(X) = \sum_{s=1}^{S} g_s(X) \cdot X_s$, where each coefficient function $g_s(X)$ is continuous, and for any $\epsilon > 0$, there exists an L-FMLC configuration (i.e., a choice of* `DeepModel` *architecture, MF parameters, and weights $W_{s,i}$) such that:*

$$\sup_{X \in D} |f(X) - y_{\textit{L-FMLC}}(X)| < \epsilon$$

Proof Sketch. The proof relies on demonstrating that the dynamically generated coefficients $C_s(X)$ of the L-FMLC model can arbitrarily approximate any set of continuous target coefficient functions $g_s(X)$. The total approximation error can then be shown to be arbitrarily small. This requires a set of standard assumptions.

Assumptions:

A.1 The input domain $X \in D$ is a compact subset of $\mathbb{R}^S$.

A.2 The base membership functions $\mu_{s,k}(\cdot)$ (e.g., Gaussian) are continuous. The overall fuzzification map from an input X to the gated membership vector $\hat{M}(X)$ is continuous. This holds as it is a composition of continuous functions (MFs, gates, etc.).

A.3 The `DeepModel` is a universal approximator for continuous vector-valued functions on a compact domain. This is a standard property for sufficiently large MLPs with appropriate non-linear activations [12].

A.4 The span of the basis functions that can be formed by the `DeepModel`'s outputs (the modulators N_i) is dense in the space of continuous functions. This follows from Assumption A.3.

The proof proceeds in a constructive manner:

1. **Approximating Target Coefficients $g_s(X)$:** For any set of continuous target coefficient functions $g_s(X)$, due to Assumption A.4, we know there exist a set of ideal modulator functions $h_i(M(X))$ and weights $\tilde{W}_{s,i}$ such that the ideal dynamic coefficient, $C_s^*(X) = \sum_i \tilde{W}_{s,i} \cdot h_i(M(X))$, can arbitrarily approximate $g_s(X)$. That is, for any $\delta_1 > 0$, we have $|g_s(X) - C_s^*(X)| < \delta_1$.

2. **Approximating Ideal Modulators with L-FMLC:** By Assumption A.3, we can configure the `DeepModel` with a sufficient number of units and layers such that its actual modulator outputs $N_i(M(X))$ can arbitrarily approximate the ideal modulator functions $h_i(M(X))$. So, for any $\delta_2 > 0$, we can ensure $|h_i(M(X)) - N_i(M(X))| < \delta_2$.

3. **Error Propagation:** The difference between the target coefficient $g_s(X)$ and the actual L-FMLC coefficient $C_s(X) = \sum_i W_{s,i} N_i(M(X))$ can be bounded by the triangle inequality:

$$\begin{aligned}
&|g_s(X) - C_s(X)| \\
&\quad \leq |g_s(X) - C_s^*(X)| \\
&\quad\quad + |C_s^*(X) - C_s(X)| \\
&\quad < \delta_1 + \left| \sum_i W_{s,i} \left(h_i(M(X)) - N_i(M(X)) \right) \right| \\
&\quad \leq \delta_1 + \sum_i |W_{s,i}| \cdot |h_i(M(X)) - N_i(M(X))| \\
&\quad < \delta_1 + \sum_i |W_{s,i}| \cdot \delta_2
\end{aligned}$$

By selecting a sufficiently expressive `DeepModel` (making δ_2 small) and appropriate basis functions h_i (making δ_1 small), the error in approximating the coefficients can be made arbitrarily small. This also holds for the learnable MF parameters (c, σ, g), as the overall mapping is continuous and the optimization process will find parameters that minimize the task loss.

4. **Bounding Total Approximation Error:** The total error between the target function $f(X)$ and the L-FMLC output $y_{\text{L-FMLC}}(X)$ is:

$$\begin{aligned} |f(X) - y_{\text{L-FMLC}}(X)| &= \left| \sum_s (g_s(X) - C_s(X)) \cdot X_s \right| \\ &\le \sum_s |g_s(X) - C_s(X)| \cdot |X_s| \end{aligned}$$

Since X is on a compact domain, $|X_s|$ is bounded. As the coefficient error $|g_s(X) - C_s(X)|$ can be made arbitrarily small, the total error can be made less than any given $\epsilon > 0$.

This demonstrates that the L-FMLC architecture is sufficiently expressive to model a wide range of complex functions. □

A.3 Convergence Analysis of L-FMLC Training

$$L_{\text{total}} = L_{\text{task}} + \lambda_{\text{sp}} L_{\text{sparse}} + \lambda_{\text{ov}} L_{\text{overlap}} + \lambda_{\text{cov}} L_{\text{coverage}} \tag{4}$$

We analyze the convergence of the training algorithm for L-FMLC, which involves minimizing the non-convex loss function L_{total} (Eq. 4) using Stochastic Gradient Descent (SGD) or its variants like Adam. Proving convergence to a global minimum for such a complex, non-convex problem is generally intractable. Instead, we show that under standard assumptions, the training algorithm is guaranteed to converge to a first-order stationary point, where the gradient of the loss function is zero.

Assumptions: Let Θ be the set of all trainable parameters in L-FMLC, including MF parameters (c, σ, g), `DeepModel` weights, and dynamic layer weights W.

- **B.1** The total loss function $L_{\text{total}}(\Theta)$ is L-smooth, meaning its gradient is Lipschitz continuous with constant L. This implies the gradient does not change arbitrarily fast: $\|\nabla L(\Theta_1) - \nabla L(\Theta_2)\| \le L\|\Theta_1 - \Theta_2\|$. This is a reasonable assumption as all components of L-FMLC (Gaussian MFs, softplus activations, common deep learning activations, regularization terms) are smooth, and compositions of smooth functions are smooth.
- **B.2** The stochastic gradients are unbiased. The gradient computed on a mini-batch is an unbiased estimator of the true gradient over the entire dataset, i.e., $\mathbb{E}[\nabla L(X_i; \Theta)] = \nabla L_{\text{total}}(\Theta)$. This holds by definition for SGD.
- **B.3** The variance of the stochastic gradients is bounded: $\mathbb{E}[\|\nabla L(X_i; \Theta) - \nabla L_{\text{total}}(\Theta)\|^2] \le \sigma^2$ for some constant σ^2. This holds if we assume per-sample gradients are bounded.

Convergence Guarantee. Under Assumptions B.1-B.3, standard results from stochastic optimization theory guarantee the convergence of SGD. Specifically, for a non-convex, L-smooth objective function, running SGD with a decaying learning rate η_t that satisfies $\sum_{t=1}^{\infty} \eta_t = \infty$ and $\sum_{t=1}^{\infty} \eta_t^2 < \infty$ (e.g., $\eta_t = 1/t$) ensures that the expected squared norm of the gradient converges to zero [13].

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\|\nabla L_{\text{total}}(\Theta_t)\|^2] = 0$$

This theoretical result ensures that the training process for L-FMLC is stable and will not diverge. The algorithm is guaranteed to find a first-order stationary point (a local minimum or a saddle point), providing confidence that the learned model parameters represent a meaningful solution to the optimization problem. While variants like Adam have more complex dynamics, their convergence properties in the non-convex setting also point to finding stationary points [14].

A.4 Experimental Setup and Model Configurations

This section provides supplementary details on the datasets, model configurations, and training protocols used in our experiments to ensure full reproducibility.

A.4.1 Dataset Details

We evaluated our framework on three distinct datasets, covering time-series forecasting, high-dimensional regression, and binary classification tasks.

- **S&P 500 Forecasting:** We used daily closing prices of the S&P 500 index [15] from October 2014 to August 2024 (2482 samples). The task is to predict the next-day price change using the five most recent daily price changes as features ($S = 5$). The data was chronologically split into a training set (70%, 1737 samples) and a testing set (30%, 745 samples).
- **Solar Irradiation Prediction:** This regression task uses data from the National Solar Radiation Database (NSRDB) [9] with 20 meteorological features ($S = 20$) to predict Global Horizontal Irradiance. The training and validation sets were drawn from 28,091 hourly samples from the year 2020 (90%/10% split). The testing set consisted of 11,269 samples from different dates across 2018, 2019, and 2021 to test for generalization.
- **Breast Cancer Wisconsin (Diagnostic):** This widely-used binary classification dataset from the UCI Repository [16] contains 569 samples with 30 features ($S = 30$). The task is to classify tumors as malignant or benign. The data was randomly split into a training set (70%, 398 samples) and a testing set (30%, 171 samples).

A.4.2 Framework and hardware

All proposed models and deep learning baselines were implemented using TensorFlow 2.14.0 in Python, with one NVIDIA T4 GPU for Solar Dataset.

A.4.3 Fuzzification and Training Protocol

Fuzzification Parameters. For the Static FMLC baseline, a fixed number of fuzzy sets (K_{fix}) was used. For our L-FMLC models, we started with a maximum number of potential sets (K_{max}) and allowed the model to prune them via our regularized gating mechanism. The specific values used are detailed in Table 3.

Table 3: Number of Fuzzy Sets (K) Used in Experiments.

Dataset	$K_{\textbf{fix}}$ **(Static FMLC)**	$K_{\textbf{max}}$ **(L-FMLC variants)**
S&P 500	5	7
Solar Irradiation	5	10
Breast Cancer	3	5

Hyperparameter Tuning and Training. A grid search was used to tune key hyperparameters for all models, including learning rates (from 10^{-5} to 10^{-2}), batch sizes (32, 64, 128), dropout rates (0.1 to 0.5), and architectural details (e.g., layers, units, attention heads). The number of modulators (n_N) for FMLC/L-FMLC was tuned over $\{1, 4, 8, 16\}$. The Adam [17] or Adamax optimizers were used with Mean Squared Error for regression and Binary Cross-Entropy for classification. Models were trained for up to 5000 epochs (S&P 500), 2000 epochs (Solar), or 100 epochs (Cancer), with early stopping based on validation loss to select the best model. For the L-FMLC regularization, λ values were searched in the range $[10^{-6}, 10^{-2}]$.

Best-Performing Model Architectures The following tables detail the final architectures for the Static FMLC baselines and our proposed L-FMLC models after hyperparameter tuning. Note that for L-FMLC, the number of modulators (n_N) was also tuned. We present the optimal configurations found.

Table 4: S&P 500 Forecasting Model Configurations (Base: LSTM).

Model Variant	DeepModel Architecture for Modulator Generation
Static FMLC	LSTM(32)-LSTM(16)-LSTM(8) (for $n_N = 8$)
L-FMLC (Ours)	LSTM(32)-LSTM(16) (for $n_N = 8$)

Table 5: Solar Irradiation Regression Model Configurations.

Base	Model Variant	DeepModel Architecture for Modulator Generation
LSTM	Static FMLC	LSTM(32)-LSTM(16) (for $n_N = 16$)
	L-FMLC (Ours)	LSTM(64)-LSTM(32) (for $n_N = 16$)
Transformer	Static FMLC	4xEnc(8h, 50u)-4xDec(8h, 50u) (for $n_N = 8$)
	L-FMLC (Ours)	4xEnc(8h, 64u)-4xDec(8h, 64u) (for $n_N = 8$)

A.5 Statistical Significance and Reproducibility

To ensure the robustness of our findings, all core models were trained and evaluated 10 times using different random seeds. This allows us to assess the statistical significance of the performance differences observed in Section 3. We report mean, standard deviation (SD), standard error of the mean (SEM), and 95% confidence intervals (CI) for the key metrics. This analysis confirms that the performance gains of our proposed L-FMLC (w/ Reg) framework are consistent and statistically significant across all benchmarks.

S&P 500 Forecasting. On the S&P 500 dataset, where all models perform well, statistical analysis is key to validating the smaller margins of improvement. As shown in Table 7, our full L-FMLC-LSTM (w/ Reg) model achieves a mean RMSE of 44.95 with a 95% CI of $[44.81, 45.09]$. The confidence interval for the Static FMLC-LSTM is $[45.53, 46.09]$. The complete lack of overlap between these two intervals provides strong evidence that our regularized adaptive approach yields a statistically significant improvement, even on a high-performing baseline. Interestingly, the CI for the Static FMLC-LSTM overlaps with that of the L-FMLC-LSTM (w/o Reg), indicating that without our proposed regularization, the adaptive model offers no reliable advantage over the simpler static framework.

Solar Irradiation Regression. The results for the Solar Irradiation dataset, a key high-dimensional benchmark, are shown in Table 8. The statistical analysis provides strong evidence for our claims. Our full L-FMLC-Transformer (w/ Reg) model achieved a mean RMSE of 91.33 with a tight 95% confidence interval of $[89.75, 92.91]$.

Crucially, this CI does not overlap with any of the other models.

- **vs. Static FMLC:**The upper bound of our model's CI (92.91) is significantly lower than the lower bound of the Static FMLC's CI (94.48), confirming that our adaptive, regularized approach is statistically significantly better.
- **vs. L-FMLC (w/o Reg):**Similarly, the CI [97.99, 101.79] for the unregularized model is clearly separated, proving that the regularization provides a statistically significant performance boost and is not just an incidental improvement.
- **vs. Standard Transformer:**The performance gap is even larger, underscoring the overall superiority of the neuro-fuzzy architecture.

Breast Cancer Classification. A similar analysis for the Breast Cancer classification task (Table 9) reinforces these conclusions. The L-FMLC-MLP (w/ Reg) achieves a mean AUC-ROC of 0.99990

Table 6: Breast Cancer Classification Model Configurations (Base: MLP).

Model Variant	DeepModel Architecture for Modulator Generation
Static FMLC	Dense(16, relu)-Dense(8, relu) (for $n_N = 8$)
L-FMLC (Ours)	Dense(32, relu)-Dense(16, relu) (for $n_N = 8$)

with a 95% CI of $[0.99982, 0.99998]$. While the absolute margins are small on this high-performing task, the confidence intervals confirm a statistically significant, albeit slight, advantage over the Static FMLC (CI: $[0.99973, 0.99987]$) and a more pronounced advantage over the unregularized L-FMLC (CI: $[0.99868, 0.99912]$). This demonstrates that even when performance is near-perfect, our regularization reliably guides the model to a more optimal and stable solution.

Table 7: Detailed Performance Metrics with Confidence Intervals for the S&P 500 Dataset.

Model	Metric	Mean	SD	SEM	95% CI Lower	95% CI Upper
L-FMLC-LSTM (w/ Reg)	R^2	0.9960	0.00018	0.000057	0.99587	0.99613
L-FMLC-LSTM (w/ Reg)	RMSE	44.95	0.20	0.0632	44.81	45.09
Static FMLC-LSTM	R^2	0.9953	0.00025	0.000079	0.99512	0.99548
Static FMLC-LSTM	RMSE	45.81	0.40	0.1265	45.53	46.09
L-FMLC-LSTM (w/o Reg)	R^2	0.9949	0.00030	0.000095	0.99468	0.99512
L-FMLC-LSTM (w/o Reg)	RMSE	46.21	0.45	0.1423	45.89	46.53
LSTM (Standard)	R^2	0.9902	0.00040	0.000126	0.98991	0.99049
LSTM (Standard)	RMSE	48.85	0.90	0.2846	48.21	49.49

Table 8: Detailed Performance Metrics with Confidence Intervals for the Solar Irradiation Dataset.

Model	Metric	Mean	SD	SEM	95% CI Lower	95% CI Upper
L-FMLC-Transformer (w/ Reg)	R^2	0.8852	0.007	0.00221	0.8802	0.8902
L-FMLC-Transformer (w/ Reg)	RMSE	91.33	2.25	0.7115	89.75	92.91
Static FMLC-Transformer	R^2	0.8621	0.008	0.00253	0.8564	0.8678
Static FMLC-Transformer	RMSE	96.27	2.50	0.7906	94.48	98.06
L-FMLC-Transformer (w/o Reg)	R^2	0.8515	0.011	0.00348	0.8436	0.8594
L-FMLC-Transformer (w/o Reg)	RMSE	99.89	2.80	0.8854	97.99	101.79
Transformer (Standard)	R^2	0.6880	0.018	0.00569	0.6751	0.7009
Transformer (Standard)	RMSE	142.32	6.00	1.8974	138.03	146.61

Table 9: AUC-ROC Statistics with Confidence Intervals for the Breast Cancer Dataset.

Model	Mean AUC	SD	SEM	95% CI Lower	95% CI Upper
L-FMLC-MLP (w/ Reg)	0.99990	0.00007	0.000022	0.99982	0.99998
Static FMLC-MLP	0.99980	0.00010	0.000032	0.99973	0.99987
L-FMLC-MLP (w/o Reg)	0.99890	0.00035	0.000111	0.99868	0.99912
MLP (Standard)	0.96650	0.00400	0.001265	0.96364	0.96936

NeurIPS Paper Checklist

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and follow the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer [Yes] , [No] , or [NA] .
- [NA] means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1–2 sentence) justification right after your answer (even for NA).

The checklist answers are an integral part of your paper submission. They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While "[Yes] " is generally preferable to "[No] ", it is perfectly acceptable to answer "[No] " provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering "[No] " or "[NA] " is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading "NeurIPS Paper Checklist"**,
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers**.

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The paper's claims in abstract and introduction are directly and accurately supported by my experimental results and analysis. The primary claim of outperforming both classical and black-box baselines is validated in Table 1, where L-FMLC demonstrates superior performance across all metrics. The claims regarding multi-level interpretability are also substantiated: Figure 2 shows the model's autonomous discovery of meaningful concepts (e.g., weather states), while Table 2 and Figure 3 illustrate the complete pipeline for distilling complex logic into actionable meta-rules and natural language reports. The paper's scope is appropriately focused on the presented solar forecasting case study, which serves as a robust proof-of-concept for the framework's value in urban systems.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.

- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper's "Discussion and Conclusion" section directly addresses limitations by identifying the need for uncertainty quantification (UQ) as a key area for future work. We acknowledge that for high-stakes urban systems, a point-estimate prediction is insufficient and that a model must also convey its own confidence to be fully trustworthy for risk-aware decisions. Additionally, by presenting the work as a detailed "case study", the paper implicitly frames the scope of its empirical claims to the solar forecasting domain, acknowledging that further validation is required for the other proposed applications.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

   Justification: This paper's primary contribution is the presentation and cognitive analysis of a framework, rather than the introduction of new theoretical proofs. While we do not present formal theorems in the main body, we briefly reference the framework's universal approximation capacity and provide a more detailed discussion of this theoretical foundation in Appendix A.2 (Universal Approximation Capacity of L-FMLC).

   Guidelines:

   - The answer NA means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.

- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Yes, the paper provides sufficient detail to reproduce our main findings. The architecture is described in Section 2. The Technical Appendix provides detailed descriptions of the datasets, key model hyperparameters, and specific implementation details in Appendix A.4 (Experimental Setup and Model Configurations). We also report on statistical significance and reproducibility measures in Appendix A.5 (Statistical Significance and Reproducibility).

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

Justification: All datasets used are publicly available, with sources and preprocessing detailed in Appendix A.4 (Experimental Setup and Model Configurations). We commit to releasing our source code on a public repository upon acceptance to further facilitate replication.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: Appendix A.4 (Experimental Setup and Model Configurations) details the datasets, including specific sources, features, train/test/validation splits, training procedures (epochs per dataset, optimizers like Adam/Adamax, loss functions), and the final architectures for all proposed models. This information is provided to ensure the results can be understood and the experimental setup is transparent.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: All predictive performance metrics reported in Section 3 are the mean results from 10 independent runs with different random seeds, as detailed in Appendix A.5 (Statistical Significance and Reproducibility).

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Appendix A.4 (Experimental Setup and Model Configurations) specifies the use of NVIDIA T4 GPUs and the TensorFlow 2.14 framework, alongside detailed model architectures and training epochs.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The work adheres to standard academic practices for responsible research.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Yes. The paper focuses on positive societal impacts, such as enabling trustworthy AI for urban sustainability. It also addresses the negative societal risk of over-reliance on the model's simplified explanations. Our framework's multi-level interpretability is presented as a direct mitigation strategy that encourages critical assessment, while the proposed future work on uncertainty quantification aims to explicitly communicate when the model's output should not be trusted.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The datasets used in this work (S&P 500 financial data, solar irradiation meteorological data, and the Breast Cancer Wisconsin (Diagnostic) clinical dataset from UCI) are standard, publicly available, or ethically sourced research datasets not typically associated with high direct misuse risks like large language models or generative AI. Our proposed L-FMLC models are specific architectures trained on these numeric datasets for regression/classification tasks and do not inherently pose a high risk for broad societal misuse upon release of the model architecture or trained instances on these specific tasks. No new, sensitive datasets were collected or scraped for this research.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: All datasets used (S&P500, NREL NSRDB for Solar Irradiation, UCI Breast Cancer Wisconsin) are publicly available and are cited appropriately in Appendix A.4 (Experimental Setup and Model Configurations). Standard open-source software libraries (TensorFlow, Scikit-learn, NumPy), mentioned in Appendix A.4 (Experimental Setup and Model Configurations), are used according to their respective licenses.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: This paper introduces a novel framework (L-FMLC) and methodologies for its training and interpretation, but it does not introduce new datasets or standalone software packages intended for release as separate "assets".

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: This paper does not involve crowdsourcing experiments or research with human subjects. The research is based on publicly available datasets (S&P 500, NREL NSRDB for Solar Irradiation, and UCI Breast Cancer Wisconsin) and algorithmic development. Therefore, participant instructions, screenshots, and compensation details are not applicable to this work.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.

- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: This research does not involve human study participants.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [NA]

    Justification: The core methodology of this research, the L-FMLC framework, does not involve the use of Large Language Models (LLMs). The proposed methods for integrating fuzzy logic with deep learning models for numerical data, including the rule extraction and clustering techniques, are developed independently of LLM technology. Any LLM usage was confined to assisting with text generation, editing, and refinement of the manuscript, not for developing or implementing the core L-FMLC algorithms or experiments.

    Guidelines:

    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.