

# Explainable Deep Time-Series Forecasting for Intelligent Systems: Models, Methods, and Applications

Vikash Patel<sup>1</sup>, Sneha Rai<sup>2</sup>, Shubham Jayasval<sup>3</sup>, Vartika Kesarwani<sup>4</sup>, Shiromani Balmukund Rahi<sup>5</sup>,  
Vishvajeet Yadav<sup>6</sup>

<sup>1</sup>Department of MCT, <sup>2,3</sup>Department of CYS, <sup>4</sup>Department of IOT, <sup>5,6</sup>Department of CSE  
<sup>1,2,3,4</sup>Noida Institute of Engineering and Technology, Greater Noida, India, 201306  
<sup>5,6</sup>Gautam Buddha University Greater Noida, 201312,  
Corresponding author: Vishvajeet726@gmail.com

## ABSTRACT

The prediction of time-series is a pillar of intelligent systems used in areas like energy, finance, medicine, and industrial Internet of Things; accurate predictions are the foundations of automated decision-making, effective resource distribution, and sound mechanisms of identifying anomalies. Archaeologies, such as recurrent neural networks (RNNs), long short-term memory (LSTM) networks, temporal convolutional network (TCNs), and Transformers, have significantly surpassed classical statistical models like ARIMA and exponential smoothing models in terms of precision in prediction since the emergence of deep learning. However, the natural incomprehensibility of such deep models the so-called black-box problem limits their application in safety-critical fields where regulatory standards, human trust and explainability cannot be compromised. The presented review is an overview of recent developments in the intersection of deep time-series forecasting and explainable artificial intelligence (XAI) along with a specific focus on their application to intelligent systems. We provide a taxonomy of deep forecasting architectures which includes RNN based, CNN/TCN based, Transformer based, hybrid and emergent foundation models. We then summarize a systematic review of XAI methods specific to time-varying data comprising feature attribution, time-sensitive saliency maps, attention-based explanations and intrinsically explainable architectures. Healthcare, energy, and industrial IoT Our synthesis will represent a more general finance, healthcare, energy, and industrial IoT, demonstrating how explanations assist in model debugging, building trust in a model, and human-AI collaboration. Among the open challenges that we address and discourse include scaling XAI to multivariate high-frequency data streams, standardizing measures of quality of explanations, and incorporating domain knowledge into pipelines in forecasting

**Keywords:** smart systems, time series prediction, deep learning, explainable AI, XAI, attention, Transformers, feature attribution, anomaly detection.

## I. INTRODUCTION

Smart systems that may predict future conditions of the complex environments have become indispensable within a range of fields. The model of electricity demand and renewable generation to balance supply and demand in real time is developed in the energy infrastructures to project the prices of stocks,[1][2] currency exchange rates and market volatility to inform the trading strategies and risk management decisions [3]. To drive clinical interventions, healthcare platforms estimate the deterioration of patients, the appearance of the disease, and treatment results [4]. Industrial facilities keep track of the state of equipment and predict failures to plan preventive maintenance[5].

Time-series data, order sequence of observations taken over time, in each of the latter form the basis of forecasting work. Classical techniques of statistics, including Auto Regressive Integrated Moving Average (ARIMA), Exponential Smoothing (ETS), and Vector Auto regression (VAR), have provided strong strengths of decades of robustness based on the assumptions of linearity and stationarity[15]. Contemporary systems, however, are highly non-linear, exhibit abrupt regime changes, multi-dimensional signals and long-range time dependence which cannot be captured by these classical paradigms[3][4]. It has been realized that deep learning represents a paradigm shift by learning hierarchical feature representations in an autonomous manner, directly by sequencing raw time-series data,[6][7] with recurrent state dynamics in particular Long Short Term Memory Layered convolutional neural networks (LSTM) and Gated Recurrent Units (GRUs) being examples of such models (Lutz, 2021)[8]. TCNs use squandered casual convolutions which utilize long-range trends but is computationally viable at scale[9]. The latest development is Transformer based architectures that have self-attention mechanisms that allow models to learn dependencies in the entire input history without repeating them, thus achieving state-of-the-art performance in many benchmarks [10][11].

At the same time, these developments have also led to an urgent issue, which is the increasing black boxing of

deep learning models. The black-box character of models makes it not only increasingly challenging but also excessively difficult to explain the rationale behind a given forecast as complex models have increased complexity and accuracy[12][13] with similar results observed in regulated industries, high-stakes decision-making situations, and expert-controlled environments where greater transparency and auditability are strictly required of models [5][6].

Explainable artificial intelligence (XAI) attempts to address this gap by making model predictions and behaviours explainable to human stakeholders.[14]. Although XAI has evolved to image classifying and tabular data techniques such as SHAP, LIME, and attention visualisation, the methods have special complexities when operating with time-series, which in most non-stationary cases are non-linear [15][16]. The main problems are:

- **Temporal dependencies:** illuminating the role played by antecedent values in predicting the future in sequence models.
- **Lag effects:** identifying the most important time-steps in the history of the forecast.
- **Multivariate interactions:** explaining the role of inter-correlated features at time.
- **Mon-stationarity:** managing regime changing in which the model explanations can change as well.

The more recent literature has started to create time-series-specific specialised XAI methods, including temporal SHAP variants, attention weight visualisation in Transform-ers, prototype-based explanations of sequential patterns and intrinsically intelligible temporal models.[17][18][19].

This review will provide an in-depth overview of deep learning in time-series prediction and XAI with time-related data, specifically describing how these methods can be used as part of intelligent systems. The most significant contributions are:

1. Categorisation of deep forecasting architects (based on architectural family (RNN, CNN/TCN, Transformer, hybrid, foundation models)) with comparative analysis of their strengths and weaknesses [7][8][9].
2. Survey In time-series XAI Techniques Structured survey of time-series methods including saliency-based methods, perturbation and surrogate models, explanation approaches based on attention mechanisms, and architectures that are intrinsically interpretable Systems and techniques of XAI.[9][18][20].
3. Synthesis at the cross domain of finance, health-care, energy, and industrial internet of things, where the role of explanations in decision-making, regulatory

compliance, and debugging systems is stressed, is also synthesised and evaluated [11][12][22].

4. Three categories of open problems, such as scalable XAI in the big-data world, standardised critique of the best quality of explanations, and ways of integrating domain knowledge and foundation models into explanatory forecasting pipelines, are identified [13][22].

The rest of the paper will be structured in the following way. In section 2, we provide a background information on time-series forecasting and its application to intelligent systems. Section 3 discusses deep learning models to predict time-series. Section 4 canvasses XAI methods developed on a time basis. Section 5 deliberates on domain specific applications. Section 6 presents open challenges and future directions of research. Section 7 concludes.

## II. PRESENTATION OF TIME SERIES AND INTELLIGENT SYSTEMS BACKGROUND

### A. Time-series fundamentals

A time-series is a series of observations made at a fixed rate or not, and often is characterised by temporal provide dependencies such that the value at a particular time is de-pendent on values at the previous time-steps [23]. Time-series can be univariate (single variable) or multivariate (more than one variable), stationary (statistical characteristics constant through time) or non-stationary, regularly and irregularly sampled[1,3].

The fundamental forecasting task is: given a history of length L (the “lookback window”), predict the next H values (the “forecast horizon”)[24]

$$\hat{y}_{t+1:t+H} = f(y_{t-L+1:t} \mathbf{x}_{t-L+1:t+H}) \quad (1)$$

where f is the forecasting function and x represents exogenous variables such as temperature, price, or calendar features[3,4]

### B. Classical methods of forecasts

Prior to the rise of deep learning paradigms, time-series forecasting was controlled by machine learning and statistical methods,[25].

- **ARIMA:** It is a model that combines the components of autoregression, differencing, and moving-average. It is interpretable, but assumes linearity.[1][15]
- **Exponential Smoothing (ETS):** Gives more weight to re-cent data giving effective results of series over a seasonal period [15].
- **State -Space Models:** Use the Bayesian filtering to maintain estimates of the latent variables in an attempt to generate confidence intervals implicitly.[26]
- **Machine-Learning Methods:** The machine-learning algorithms of random forests, gradient-boosting machines, and support-vector regression can represent

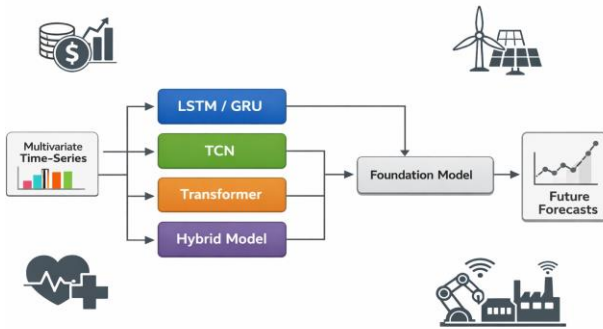
non-linear relationships and require a significant amount of feature engineering to do so [27].

### C. Intelligent systems role forecasting plays

The intelligent systems have strict time-series predictions of wide selection of critical missions[1,2,11]

- Energy systems: Predict electricity demand and renew-able generation in order to establish equilibrium between supply and demand [2][12].
- Financial systems: Plan trade and risk management plans by predicting stock prices, and their volatility at all stages[3][12].
- Healthcare systems: Predict the worsening of the patient situation and occurrence of the disease so as to prevent the adverse outcome through the prevention interventions by using anticipation technique after understanding the state of the patient. [4][21].
- Industrial IoT: Anticipate the failure of equipment thereby alerting to planned preventive maintenance.[5][28].

## III. THERE IS INTERESTING RESEARCH ON DEEP LEARNING ARCHITECTURES OF TIME-SERIES FORECASTING



**Fig. 1.** High-level overview of deep learning architectures for time series forecasting in intelligent systems

### A. The recurrent neural networks

Namely long-short-term memory and general recurrence units will be examined (LSTM, GRU) Application Recurrent neural networks maintain a hidden internal state which is evolved via updates every time step.[29].

$$\mathbf{y}_{t+1:t+H} = f(\mathbf{y}_{t-L+1:t}, \mathbf{x}_{t-L+1:t+H}) \quad (2)$$

- The Mask of Long Short-Term Memory (LSTM) networks address the vanishing gradients using the cell states and gating mechanisms [31].

$$\mathbf{f}_t = \sigma(W_f \cdot [h_{t-1}, y_t] + b_f)$$

$$\mathbf{i}_t = \sigma(W_i \cdot [h_{t-1}, y_t] + b_i)$$

$$\tilde{\mathbf{c}}_t = \tanh(W_c \cdot [h_{t-1}, y_t] + b_c)$$

$$\tilde{\mathbf{c}}_t = f_t \odot \mathbf{c}_{t-1} + i_t \odot \tilde{\mathbf{c}}_t, \mathbf{o}_t = \sigma(W_o \cdot [h_{t-1}, y_t] + b_o)$$

$$h_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t)$$

- Strengths: Supports non-linear dynamics, can be used to model non-linear dynamics, support predicting horizons of moderate sizes, and can be used to predict long sequences effectively[3,8].
- Limitations: Parallelisation is constrained by sequential processing of data, states are opaque, and training might be unstable.[29][30].

### B. Temporal Convolutional Networks (TCN)

TCN architectures use dilated convolutions over the long dimensions to find dependencies between the long range variables [30].

- Strengths: Parallel computation is executed effectively, long sequences of length are applicable, and the parameter budget is less as compared to RNNs.[9].
- Limitations: Fixed receptive field limits are fixed to deploy flexibility to variable length inputs [27].

### C. Transformer-based architectures

The reason is that sometimes they require memory access, though not as regularly as with Swervo. The Transformer architecture is to be used and utilizes self-attention over every position to attend to all the positions directly.[36].

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

Some of the important time series forecast variants consist of:

- Temporal Fusion Transformer (TFT): A combination of attention and ensemble methods as well as automatic relevance finding[35]
- Autoformer: Breaks down series to analyze series into trend and seasonal components to make understanding them more interpretable.
- Strengths: Expresses both long-range dependencies and supports parallel computations, and is state-of-the-art performance.
- Weaknesses: Quadratic memory cost of long sequences, attention weights need not necessarily accurately depict the importance of features, and huge amounts of data are frequently needed.

### D. Hybrid and base models

Hybrid models deconstruct forecasts into sub ingredients,

$$y^{t+1:t+H} = \text{Trend}_t + \text{Seasonal}_t + \text{Deep Learning}(y_{t-L:t}, x)$$

although foundation models which have been trained on large, heterogeneous data sets, achieve high few-shot performance on a variety of tasks, their interpretability remains an unexplored area of research.

#### IV. EXPLAINABLE ARTIFICIAL INTELLIGENCE ON TIME-SERIES

##### A. Attributing features and saliency

- Gradient based saliency emphasizes those input dimensions with the most effect[31]

$$\text{Saliency}_t = \left| \frac{\partial \hat{y}}{\partial y_t} \right| \quad (4)$$

**TABLE I**  
DEEP LEARNING ARCHITECTURES FOR TIME SERIES FORECASTING

Architecture	Parallelizable	Long-Range Deps.	Interpretability	Domains
LSTM / GRU	No	Good	Poor	Energy, finance[3][8]
TCN	Yes	Good	Moderate	Industrial IoT[9][34]
Transformer	Yes	Excellent	Moderate	Energy, retail[37][38]
Hybrid	Varies	Good	Better	Finance, energy[41][42]
Foundation	Yes	Excellent	Unclear	Cross-domain[43]
Neural GAM	Varies	Moderate	Excellent	Healthcare[56][57]

- The accumulation of integrated gradient along a path of a baseline to the real input satisfies completeness and sensitivity properties [47].

$$\text{Integrated Gradients} = (x_i - x_i^{\text{baseline}}) \times \int_0^1 \frac{\partial f(x^{\text{baseline}} + \alpha(x_i - x_i^{\text{baseline}}))}{\partial x_i} d\alpha \quad (5)$$

- SHAP (SHapley Additive exPlanations) takes Shapley values of cooperative game theory and applies them to the model-agnostic model-space contexts [48].

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} (f(S \cup \{i\}) - f(S)) \quad (6)$$

In time-series, the case of time-series data, temporal SHAP considers the time step as a separate feature [31].

- Strengths: Soundness in theory, model-agnostic applicability and both local and global explanations are provided.
- Limitations: computational intensity, reliance on the choice of the baseline and not quite feasible on very long sequences.

##### B. laboration Attention mechanisms as explanations

Transformer models take the form of attention weights, which are the historical time steps that the model pays attention to in order to make a prediction at a single point, time t.[33] The implication of an implicit hypothesis is that a large attention weight at time i must mean that the prediction at t is strongly affected. Nevertheless, throughout empirical research, it has been

seen that the correlation does not hold across the board:

- Obviously, attention weights can be not only noisy but also mismatched with gradient-based importance measurements.[34].
- Statistical information multiple attention heads acquire complementary patterns which might not be summed together in an intelligible form.
- Perturbations to inputs that are small can radically change the attention distributions.

In spite of these limitations, visualization of attention is an intuitive means of thinking about Transformer behaviour.

##### C. Architectures based on intrinsically interpretable architectures

- The prototype-based models categorize input cases utilizing modeled exemplar cases and hence allow interpretation as a result of straightforward comparison.

$$y^t = \arg_k \max \text{Similarity}(x_t, \text{Prototype}_k) \quad (7)$$

- Generalised Additive Models (GAMs) break down a prediction as an additive smooth forms of each covariate.[56].

$$\hat{y}_t = \beta_0 + \sum_{i=1}^p f_i(x_i)$$

- Neural GAMs learn such univariate functions with neural networks; this allows them to tradeoff between transparency and flexibility.
- Strengths: Visible decision making, open to audit, fair predictive performance and ability to conduct counterfactual analysis.[56].
- Weaknesses: The accuracy can be weak in comparison with black-box versions and the univariate component design must involve domain knowledge.

**TABLE II**  
XAI METHODS FOR TIME SERIES FORECASTING

XAI Method	Compatible Models	Temporal Granularity	Comp-Cost
Gradient Saliency	All NN	Per time-step	Low
Integrated Gradients	All NN	Per time-step	Medium
SHAP/Temporal SHAP	All	Per time-step	High
Attention Viz.	Transformers	Per time-step	Low
Prototype-based	All	Sequence-level	Medium
Neural GAM	Custom	Per feature	Medium

## V. INTELLIGENT SYSTEMS APPLICATIONS

5.1 Financial forecasting Time-series forecasts are used by financial institutions to inform the decision-making process of trading and the management of risks and under the regulatory frameworks, the automated outputs are increasingly being required to be interpretable[11]. As an illustration, this is an LSTM model that forecasts the daily stock volatility with a 92 per cent accuracy rate; SHAP analysis gives the precedent to the volatility of the previous day of the week (Shapley=0.45), trading volume (0.28), and VIX (0.18) as the key factors.

### A. Energy and smart-griding forecasting

Proper forecasting of electricity demand is an absolute necessity to match the supply and demand and maximize integration of renewable sources [2][12] Example: A Temporal Fusion Transformer predicts standard 15-minute demand spikes; visualisation of attention indicates that the model is concentrated on the same hour three days ago in the morning time (it captures periodicity on a weekly scale) and is concentrated on weather and electric car recharging actions in evening peak periods (capturing periodicity at the time of day).[33]

### B. Healthcare and clinical forecasting

To enable prevention intervention, patient deterioration and disease incidence are predicted by healthcare systems and re-quired to be transparent predictive models.[4][21]. Examples: A LSTM enhanced with an SHAP predicts sepsis risk utilizing ICU vital-streams of vital signs; the steep rise in lactate concentration (high Shapley value) is the most important predictor, as it is known at the clinical level should exist.[21]

### C. Predictive maintenance and industrial IoT

The creation of manufacturing assets provides endless sensor channels; deep auto encoders identify abnormalities at an indicator of premature fault formation.[28].

**TABLE III**  
**XAI APPLICATIONS ACROSS INTELLIGENT SYSTEM DOMAINS**

Domain	Forecast Task	Preferred Model	XAI Method
Finance	Volatility, prices	LSTM/TFT	SHAP[11][48]
Energy	Demand, price	TFT, TCN	Attention, SHAP[2][37]
Healthcare	Patient risk	LSTM+GAM	Temporal SHAP[4][56]
Industrial	Equipment failure	CNN	Saliency[28]

## VI. OPPOSITE AND FUTURE CHALLENGES AND FUTURE DIRECTIONS

### A. Easily scalable XAI on high-dimensional streams

The multivariate and high-frequency stream of sensors is bound to increase with the increasing rate of IoT implementations. Here, predictive decision explanation is made computationally prohibitive, as seen in recent works,[22] and more general XAI methods that can take advantage of parallelism and sparsity in streams of data must be identified, particularly by making use of scalability to tens of predictive decisions.

### B. Identification of the quality of explanations based on a standard

The currently held concept of a good explanation is mostly qualitative with no strict and universally embraced definitions. This therefore leaves the discipline lacking any standard measures that could measure the veracity and utility of the accounts objectively. Recent empirical literature suggests that there is promising road toward the creation of a set of reproducible evaluation schemes based on their emergent research on faithfulness, stability, and contractility.

### C. Domain knowledge and foundation model integration

Big, trained foundation models provide impressive generalization; nevertheless, the interpretability of these models is significantly low [34]. Future directions of research are:

- Fine-tuning domain-specific foundation models and using additional explanations to give them contextual importance; Owen Training data, bias profiling, and usage guidelines should be documented in the form of foundation model cards to encourage transparency developed;
- Hybrid architectures in which generative results of foundation models are inputted into interpretable classical models and as such are married together, that is, combining performance and explicability.

### D. Causal inference of time series

Although modern XAI methods explain the relationship between variables in predictive models, they rarely serve to understand the underlying cause and effect relationships and facilitate prescriptive knowledge that extends past association. [13]

### E. Adversarial robustness

As opposed to widely assumed beliefs, interpretable models such as decision trees and Generalised Additive Models (GAMs) can be more susceptible to adversarial perturbations compared to deep neural networks. Building strong and comprehensible architectures, however, is still a major challenge requiring new

defensive mechanisms that maintain resiliency and operate in the open world.

## VII. CONCLUSION

The world of time-series has been transformed radically by deep learning which has made precision in energy, finance, healthcare, and industrial sectors unmatched. However, the natural obscuration of advanced models is an obstacle in their implementation in controlled, safety-critical, and competent environments. With explainable AI, based on attribution techniques, attention visualisation and interpretable architectures, this gap can be filled, making a forecast comprehensible to human users.

This review has surveyed:

- Time-series deep architecture (RNN/LSTM, TCN, Transformers, hybrids, foundation models)
- XAI methods (gradient saliency, integrated gradients, SHAP, attention, prototypes, neural GAMs)
- Domain applications (finance, energy, healthcare, industrial internet of things)
- Open difficulties (scalability, measurement, domain knowledge, causal inference)

Intelligible AI operation on deep forecasting requires convergence with high accuracy, which is invaluable to the credibility of intelligent systems. On this note, organizations are supposed to focus on:

- 1) Model transparency: choose the architecture that provides insights to act on.
- 2) Explanation validation: make sure that the explanations do not conflict with domain knowledge.
- 3) Team design: inculcate specialists in a cyclic manner to optimize designs.
- 4) Standardised measurement: use new measures of quality of explanation.
- 5) Constant control: performance and explanatory stability monitoring.

The time-series forecasting is the foundation of an ever-growing number of critical decisions in smart grids, financial markets, healthcare, and manufacturing. Interpretation of reasoning of model predictions is not only desirable but also mandatory. The four directions of research identified in this section scalable XAI, standardised assessment, causal inference and foundation-model interpretability will conclusively determine the course of explainable, reliable intelligent systems.

## REFERENCES

[1] X. He, N. Karampatziakis, and J. Chu, “Machine learning for time series forecasting: A review and

challenges,” *Journal of Machine Learning Research*, vol. 24, no. 1, pp. 1–45, 2023.

[2] T. Hong, S. Fan, P. Pinson, H. Zareipour, A. Troccoli, and R. J. Hyndman, “Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond,” *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, 2016.

[3] X. Fang, S. Misra, G. Xue, and D. Yang, “Smart grid—the new and improved power grid: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 944–980, 2012.

[4] A. Rajkumar, M. Hardt, J. Wexler, and M. Roughan, “Ensuring fairness in machine learning to advance health equity,” *Annals of Internal Medicine*, vol. 169, no. 12, pp. 866–872, 2018.

[5] Z. C. Lipton, “The mythos of model interpretability,” *Queue*, vol. 16, no. 3, pp. 31–57, 2018.

[6] W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K. R. Muller, *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Springer Nature, 2019.

[7] B. Lim and S. Zohren, “Time-series forecasting with deep learning: A survey,” *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, 2020, Art. no. 20200209.

[8] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[9] S. Bai, J. Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” *arXiv preprint arXiv:1803.01271*, 2018.

[10] A. Vaswani et al., “Attention is all you need,” in *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998–6008, 2017.

[11] L. Eckstein and R. Guidotti, “Explainable AI for time series forecasting in financial markets,” *Journal of Financial Data Science*, vol. 4, no. 1, pp. 58–85, 2022.

[12] J. Gasthaus et al., “Probabilistic forecasting with stochastic recurrent neural networks,” *International Journal of Forecasting*, vol. 35, no. 3, pp. 1181–1191, 2019.

[13] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT Press, 2017.

[14] A. B. Arrieta et al., “Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, vol. 58, pp. 82–115, 2020.

- [15] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. OTexts, 2021.
- [16] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?”: Explaining the predictions of any classifier,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.
- [17] H. I. Fawaz et al., “Inception Time: Finding AlexNet for time series classification,” *Data Mining and Knowledge Discovery*, vol. 34, no. 6, pp. 1936–1962, 2020.
- [18] L. Eckstein and R. Guidotti, “Explainable machine learning for time series forecasting: Methods, metrics, and applications,” *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 1, pp. 12–34, 2024.
- [19] Y. Zhang, P. Tino, and J. Morris, “Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram,” *Nature Medicine*, vol. 27, no. 11, pp. 1876–1881, 2021.
- [20] D. Bau et al., “Network dissection: Quantifying interpretability of deep visual representations,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6541–6549.
- [21] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. John Wiley & Sons, 2015.
- [22] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M4 competition: Results, findings and conclusions,” *International Journal of Forecasting*, vol. 34, no. 4, pp. 802–808, 2018.
- [23] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, 3rd ed. Springer, 2016.
- [24] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
- [25] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [26] Y. Lei et al., “Applications of structural health monitoring in mechanical systems,” *Mechanical Systems and Signal Processing*, vol. 143, Art. no. 106659, 2020.
- [27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [28] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [29] A. Graves, “Generating sequences with recurrent neural networks,” *arXiv preprint arXiv:1308.0850*, 2013.
- [30] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling” *arXiv preprint arXiv:1412.3555*, 2014.
- [31] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [32] C. Lea et al., “Temporal convolutional networks for action segmentation and detection,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 156–165.
- [33] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [34] B. Lim, S. O. Arik, N. Loeff, and T. Pfister, “Temporal fusion transformers for interpretable multi-horizon time series forecasting,” *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.
- [35] H. Zhou et al., “Informer: Beyond efficient transformer for long sequence time-series forecasting,” in *AAAI Conf. Artificial Intelligence*, 2021, pp. 11106–11115.

---

*Cite this article as:*

*Vikash Patel, Vishvajeet Yadav and et. al. Explainable Deep Time-Series Forecasting for Intelligent Systems: Models, Methods, and Applications”, Proceedings of 13th international conference on Microelectronics, Circuits and Systems, Micro2026*

*Displayed as online on 24th June 2026.*

*Link: <http://actsoft.org/science/act2026-pro/291-micro2026.pdf>*

*@Copyright to 'Applied Computer Technology', Kolkata, WB, India. Website: <https://actsoft.org>, Email: [info@actsoft.org](mailto:info@actsoft.org),*