Research On Complex Sequence Prediction and Interpretation Based on Adversarial Learning Over Precise Networks and Heuristic Optimization Optimization

Yuhe Pan *
School of Information, Harbin GuangSha College, Harbin, China, 150025
* Corresponding Author Email: 17745609063@163.com

Abstract. In the field of high-dimensional dynamic system modeling, achieving high-precision prediction of target variables, interpretability of model structures, and quantification of output uncertainty has always been a key focus of intelligent system research. Traditional methods encounter bottlenecks in terms of feature heterogeneity, multi-dimensional interactions, and prediction credibility. This paper proposes a multi-objective sequence modeling framework that integrates generative adversarial learning, simulated annealing optimization, and a game-theoretic interpretation mechanism. This framework first cleans, standardizes, and imputes missing values in the original structured data, and then reconstructs the input tensor using a BP neural network. Subsequently, a prediction model based on GAN is constructed, and SA is used for global optimization of key hyperparameters. During the prediction phase, a Monte Carlo perturbation strategy is introduced to quantify uncertainty. Finally, the importance of features is analyzed based on the SHAP method to achieve causal transparency. This method outperforms mainstream baseline models such as LSTM and Transformer in multiple metrics, demonstrating good generalization ability and interpretability. It is suitable for prediction and auxiliary decision-making in complex engineering systems in fields such as finance and industry, and has broad application prospects and significant practical value.

Keywords: Generative Adversarial Networks, Sequence Modeling, Simulated Annealing, Model Interpretability, Feature Contribution Degree.

1. Introduction

In practical engineering, tasks such as resource scheduling and supply chain optimization rely on accurate prediction of key indicators in dynamic systems and identification of sensitive factors. Engineering systems are complex, with strong temporal correlations among variables, indicating significant time-series dependencies. System states evolve over time in a patterned and interconnected manner. The feature space is high-dimensional and heterogeneous, encompassing structured data and information from multiple heterogeneous sources. However, observed data is often incomplete, with issues like missing values, noise interference, and delayed sampling, complicating system analysis and decision-making.

Traditional complex sequence prediction methods, mainly regression models or black-box deep networks (e.g., LSTM,Transformer, show limitations in real-world applications. These models lack effective mechanisms for modeling dynamic evolutionary structures, making it hard to discern underlying patterns and dynamic changes in complex sequence data. This restricts their generalization capabilities, leading to poor performance on new datasets. Moreover, they are highly sensitive to hyperparameters, relying heavily on extensive and meticulous parameter tuning. Their inability to automatically adapt to diverse data characteristics and prediction tasks increases operational complexity and cost for practical deployment. Additionally, due to their internal complexity and black-box nature, these models struggle to output clear, interpretable prediction results, failing to explain why predictions are made and limiting their use in scenarios requiring interpretable decision-making.

In sequence modeling and optimization research, scholars have proposed innovative methods. M. A. Syed et al.[1] introduced an NNMPC model using an EAS-based neural network to generate control signals, simplifying the modeling of controlled objects in complex EAS systems for prediction. Xinran Liu et al.[2] proposed a VMD - PCA - LSTM combined model, using VMD to capture data changes and reduce PV data randomness, PCA for dimensionality reduction, and LSTM for dynamic modeling and prediction. Wei Huo et al.[3] employed parallel prediction and optimization, using Optuna for hyperparameter optimization on the XGBoost model. Julio C. Antunez et al.[4] used the Elastic Net model, combining L1 and L2 regularization, which is advantageous in handling high-dimensional data with strong feature correlations and redundancy. Z. Wang et al.[5] introduced the PRO-TIME framework driven by multimodal learning, integrating netlist and layout information and using customized models to improve performance.

To address these challenges, this paper designs a modular prediction framework called PRECISE-Net, integrating generative adversarial optimization and interpretation mechanisms. It combines three key technical modules: Generative Adversarial Learning (GAN)[6], which introduces a generator-discriminator mechanism for dynamic game optimization between sample generation and structure discrimination, better fitting non-linear prediction targets and improving generalization ability under multi-dimensional complex inputs; Simulated Annealing Optimization (SA)[7], used for global heuristic search optimization of the GAN model's structural parameters and learning strategies to avoid local optima; and Interpretation Mechanism (SHAP)[8], which introduces Shapley value theory from game theory to measure the marginal contribution of each causal variable to the model output, achieving both prediction and interpretation.

The proposed method addresses critical issues such as poor model adaptability, inadequate prediction stability, and lack of transparency in complex prediction tasks, demonstrating notable advantages in deployability, prediction accuracy, model stability, and interpretability, laying a solid foundation for its promotion in practical applications.

2. Related Work

2.1. Research Status of Sequence Modeling, Optimization, and Interpretation Methods

In complex sequence prediction, advancements in sequence modeling, model optimization, and interpretability have been accompanied by persistent challenges.

Sequence modeling has evolved with the advent of deep learning architectures. LSTM networks excel in medium-length time-series prediction tasks due to their ability to capture long-term dependencies. GRU variants enhance training efficiency by reducing parameter complexity. The Transformer architecture, incorporating attention mechanisms, has demonstrated robust performance across natural language processing, financial forecasting, and power dispatching. However, current models face limitations, particularly in handling high-dimensional heterogeneous data. Traditional RNN structures struggle to model complex nonlinear interactions among features, and many models lack interpretability, treating prediction mechanisms as black boxes.

Model optimization strategies, particularly hyperparameter tuning, are critical for performance. Grid search, while comprehensive, is computationally expensive and prone to the curse of dimensionality as hyperparameter spaces grow. Bayesian optimization improves efficiency by using probabilistic models to guide searches, yet it risks converging to local optima. Heuristic methods, such as genetic algorithms and particle swarm optimization, suffer from poor interpretability and inconsistent convergence, leading to variability in results across different runs. A critical challenge remains in identifying optimal model architectures and hyperparameters without sacrificing computational efficiency.

Interpretability has gained prominence, especially as AI models are deployed in critical domains like medical diagnosis, financial risk assessment, and autonomous driving[9]. The SHAP method, rooted in Shapley value theory, quantifies the marginal contribution of input features to model outputs, supporting model credibility evaluation and decision visualization. Despite its utility, SHAP and similar methods are constrained by assumptions of static input data, limiting their effectiveness in dynamic generative models where internal feature interactions evolve. This constraint restricts their applicability in complex dynamic scenarios.

2.2. Integrated Innovation Points of This Study

To address the trade-off challenge of existing sequence modeling methods in balancing prediction accuracy, model robustness, and structural interpretability, this study proposes an innovative integrated closed-loop modeling framework, PRECISE-Net, to simultaneously enhance multiple performance objectives. This framework deeply integrates Generative Adversarial Networks (GAN), Simulated Annealing (SA) algorithm, and SHAP analysis mechanism to construct a complete technical chain covering data modeling, strategy optimization, and result interpretation, providing an integrated solution for complex prediction tasks.

In the prediction modeling phase, considering the advantages of GAN in data distribution fitting and nonlinear feature extraction, it is introduced into model construction to enhance the ability to depict the inherent laws of complex data, laying a data foundation for subsequent steps.

In the optimization strategy phase, aiming at the global optimization problem in non-convex spaces, SA is adopted to optimize model parameters, enabling it to escape local optima, improve the model's adaptability and stability in multiple scenarios, and enhance robustness.

In the model interpretability phase, SHAP analysis mechanism is introduced to measure the factor contributions of model output results, quantify the marginal contributions of features, provide intuitive and transparent explanations for model decisions, enhance credibility and acceptability, and also provide guidance for feature engineering optimization.

This study integrates adversarial learning, heuristic hyperparameter optimization, and factor interpretation mechanisms to construct a unified and efficient closed-loop modeling system, achieving a transition from solely pursuing prediction accuracy to considering both decision transparency and model robustness, providing new paths and references for research and engineering applications in related fields, with significant theoretical value and application prospects.

3. Methodology

3.1. Overview of the Modular Prediction Framework Integrating Generative Adversarial Optimization and Interpretation Mechanisms

As shown in Fig. 1, this study constructs an integrated "prediction - optimization - interpretation" trinity modeling framework PRECISE-Net. It systematically integrates a data preprocessing module, a core prediction module, a hyperparameter optimization module, and an interpretation module, forming a full-process modeling system covering data processing, model construction, performance optimization, and result interpretation. Specifically, the data preprocessing module targets multi-source structured time-series data. It efficiently handles key steps like data cleaning, feature transformation, and feature engineering to ensure high-quality and highly available data input for subsequent model training.

The core prediction module is built on the Generative Adversarial Network (GAN). It leverages GAN's excellent ability in data distribution fitting and high-dimensional complex feature modeling to deeply model and accurately predict the dynamic evolution patterns in time-series data, effectively enhancing the model's ability to depict nonlinear time-series relationships.

The hyperparameter optimization module takes the Simulated Annealing (SA) algorithm as its core. It uses a heuristic global search strategy to efficiently optimize model hyperparameters, overcoming the issue of easily getting stuck in local optima in non-convex optimization spaces, thus significantly improving the model's generalization ability and robustness in complex scenarios.

The interpretation module introduces the SHAP (SHapley Additive exPlanations) analysis mechanism. By quantifying the marginal contribution of each input feature to the model's prediction results, it reveals the key factors affecting the predictions. This not only provides a clear and intuitive explanation basis for the model's decision-making process, enhancing the model's transparency and credibility, but also offers strong support for the optimization of feature engineering and the iterative improvement of model performance.

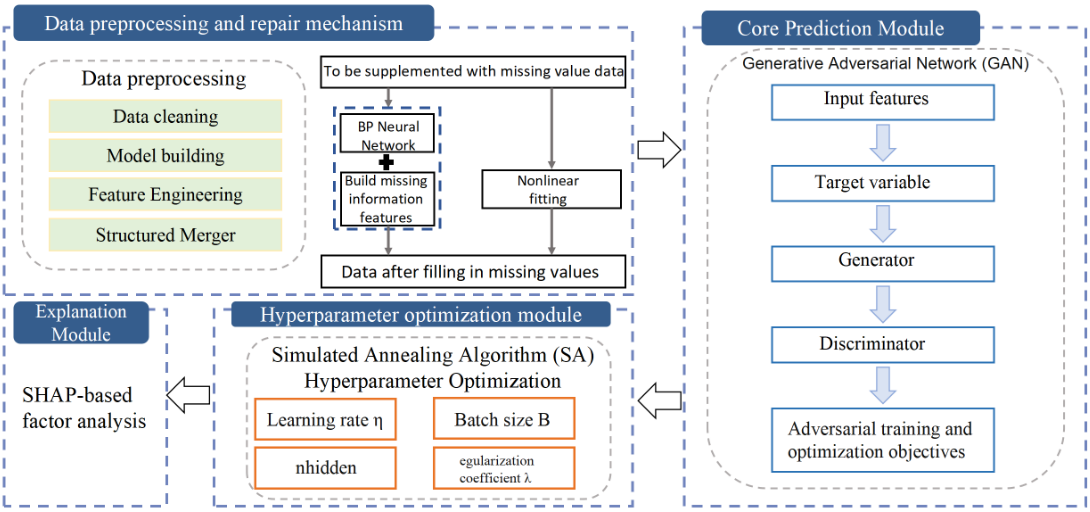

Figure 1. PRECISE-Net: A Modular Prediction Framework Integrating Generative Adversarial Optimization and Interpretation Mechanisms

In summary, this integrated architecture can achieve end-to-end efficient modeling, accurate prediction, and in-depth interpretable analysis of multi-source structured time-series data. It also possesses high embeddability, good scalability, and strong adaptability, allowing it to be flexibly applied to various complex time-series analysis tasks and providing solid technical support and theoretical basis for scientific decision-making in complex systems.

3.2. Data Preprocessing and Repair Mechanisms

In data preprocessing, this study merges structural features of different dimensions (e.g., time, categories, continuous variables) in a unified input space via structured operations to create a more informative feature set. For non-uniform sampling and missing data, BP neural networks are used for nonlinear imputation, estimating missing values accurately. Additionally, derived features like interaction terms and periodic indicators are generated to enrich model inputs. These preprocessing steps yield a high-quality, complete input tensor, aiding in model training, inference, and performance improvement.

3.3. Sequence Prediction Module: A Generative Prediction Model Based on GAN

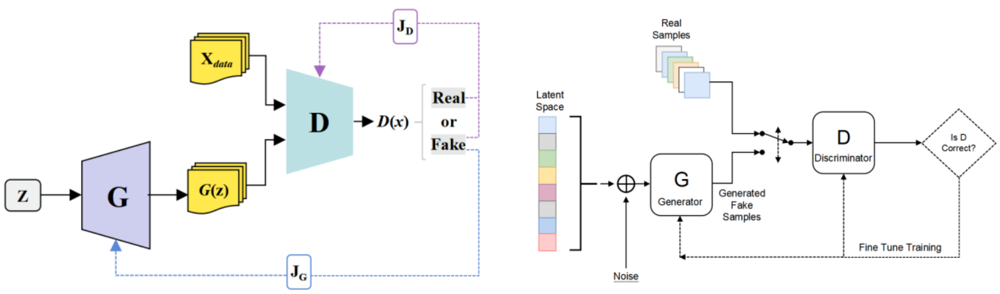

Figure 2. Model Architecture Diagram of Generative Adversarial Network (GAN)

As Fig. 2 shows, in the proposed architecture, generator G (·) simulates real system evolution logic by deep-learning historical input features to generate predicted sequences of target variables, depicting the system's future state. Discriminator D (·) acts as a "supervisor", judging the authenticity of generated sequences and driving the generator to optimize via feedback, making its sequence distribution approach real data and improving prediction accuracy. The objective function has two parts (as shown in formulas 1 and 2), aiming to maximize the generator-discriminator game for model convergence and performance.

① Generator's goal: Minimize the discriminator's ability to discriminate against the generated data, that is, maximize the discriminator's misjudgment probability of the generated data.

$$L_G = -E_{X \sim p_{data(X)}}[LogD(G(X))] \tag{1}$$

Where G represents the non-linear mapping function of the generator, and usually adopts LSTM, GRU, or a fully connected network structure.

② Discriminator's goal: Maximize the ability to distinguish between real data and generated data.

$$L_G = -E_{X \sim p_{data(Y)}}[Log\sigma(\mathrm{f}(\mathrm{Y}; \theta\mathrm{d}))] - E_{X \sim p_{data(X)}}[Log(1 - D(G(X)))] \tag{2}$$

Where f represents the non-linear mapping function of the discriminator (such as a fully connected network), σ is the Sigmoid activation function, and θd is the trainable parameter of the discriminator.

The proposed architecture supports multi-step prediction, forecasting both specific future values and sequences, aiding decision-making. It integrates attention mechanisms to model long-range dependencies, enhancing handling of complex time-series data. Its flexible structure allows substitution of generator and discriminator with LSTM or Transformer variants to suit different performance and efficiency needs. Compared to traditional models, it captures complex nonlinear structures and cross-dimensional interactions, crucial in finance and meteorology. It also addresses boundary prediction weaknesses, achieving high accuracy in edge regions and broadening applicability.

3.4. Adaptive Hyperparameter Optimization: Simulated Annealing (SA)

At the algorithm design level, simulated annealing is a heuristic global optimal search strategy, with its core principle being the simulation of the solid annealing process. As Fig. 3 shows, at the initial search stage, it allows accepting "suboptimal solutions" by taking non-optimal ones with a certain probability, effectively avoiding the algorithm getting stuck in local minima and enhancing the chance of finding the global optimum. In this study, it's used to optimize model hyperparameters, including learning rate, hidden layer units, time window length, regularization strength, and batch size. Optimizing them and updating the objective (loss function on the validation set as in Formula 3) can boost the model's generalization and prediction accuracy.

$$\mathrm{L_{val}} = \frac{1}{\mathrm{N_{val}}} \sum_{\mathrm{i}=1}^{\mathrm{N_{val}}} ||\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}||_2^2 \tag{3}$$

Where is the true value, is the predicted value, and is the number of samples in the validation set.

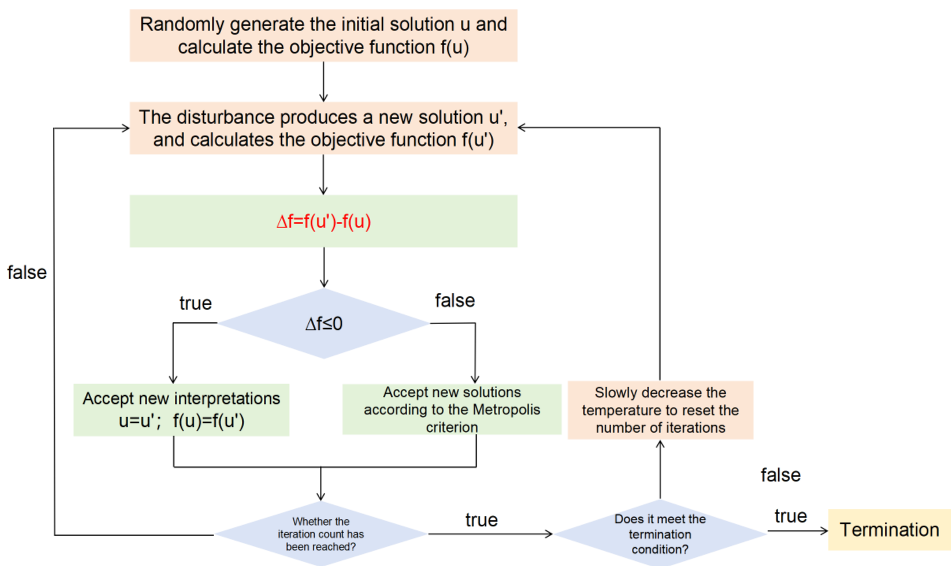

Figure 3. Flowchart of Simulated Annealing Hyperparameter Tuning

During optimization, the validation set loss function or prediction performance metrics (such as MSE, R^2) are used as the objective functions to search for the global optimal solution. The validation set loss function reflects the model's fitting ability to unseen data, while the prediction performance metrics measure the prediction accuracy. By guiding the hyperparameter optimization with these, the model's performance on the validation set and its reliability in practical applications can be improved. Compared with traditional hyperparameter tuning methods, the simulated annealing strategy has significant advantages. Its probabilistic acceptance mechanism makes the search process more stable, avoiding performance fluctuations, and overcomes the problems of traditional methods which rely on manual experience, have low efficiency, and produce unstable results. Moreover, it does not rely on specific model structures and can be connected to any black-box model, demonstrating strong universality and being widely applicable to various machine learning tasks.

3.5. Interpretable Mechanism: Factor Analysis Based on SHAP

In the model interpretability analysis phase, this study adopts the SHAP (SHapley Additive exPlanations) method for in-depth exploration. The SHAP method is rooted in the theoretical framework of cooperative game theory, and its core mechanism is to assign a stable interpretation weight to each input variable by calculating the marginal contribution feature by feature. Specifically, this method decomposes the changes in the model's prediction results into the sum of the marginal contributions when each input feature acts alone, thus achieving a quantitative evaluation of the importance degree of each feature and providing a powerful tool for understanding the model's decision-making process.

For a given feature and model output, the Shapley value represents the contribution of feature to the model's prediction result, and the calculation formula is as follows:

$$\phi(f) = \sum_{S\subseteq N\setminus\{f\}} \frac{|S|!(|N|-|S|-1)!}{|N|!}[f(S \cup \{f\}) - f(S)] \tag{4}$$

Where is the feature set, is the set of all features, is the result of model prediction using only the feature set, and is the prediction result after adding feature to the feature set. The Shapley value is calculated by traversing all possible feature combinations to fairly allocate the contribution of each feature to the prediction result.

The SHAP method has significant value in applications. In terms of ranking factor influences, calculating the SHAP values of features can accurately identify the driving factors that play a key role in the prediction results, providing support for feature selection and engineering optimization. In the recognition of feature interaction patterns, it can unveil the interrelationships among variables, facilitating the comprehension of the intricate data mechanisms and serving as a basis for constructing precise prediction models. In addition, it enhances the transparency of the model through visualizing the decision-making mechanism, making the decision-making process intuitive and easy to understand, and providing strong support for strategy formulation.

4. Experiment Design & Results

4.1. Experimental Objectives and Settings

The objective of this study is to verify the effectiveness and stability of the proposed "prediction - optimization - interpretation" integrated modeling framework in complex sequence prediction tasks. This is achieved by examining the framework's performance in terms of prediction accuracy and generalization ability, hyperparameter optimization effectiveness, uncertainty expression capability, and interpretability of factor contributions. In terms of data composition, the input dimensions include historical observational data, category identification information, and structured statistical features, etc. The output target is a multivariate prediction sequence over a continuous time period. The dataset is divided into a training set (accounting for 70%), a validation set (accounting for 15%), and a test set (accounting for 15%) based on chronological order.

4.2. Performance Evaluation Metrics

To comprehensively and systematically evaluate the model's prediction performance, this study adopts the following evaluation criteria (as shown in Equations 5-8).

The coefficient of determination (R^2) is an important indicator for measuring the model's fitting ability to the true values. Given a dataset containing n samples, with the true values denoted as (i=1,2, ···,n), the model's predicted values denoted as $\hat{\mathbf{y}}_{\mathbf{i}}$(i=1,2,···,n), and the mean of the true values denoted as $\bar{y}$.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

Its value ranges from negative infinity to 1, and the closer it is to 1, the better the model fits the data, indicating that the model can more accurately capture the inherent patterns and trends in the data.

For the uncertainty analysis of predicted values, Monte Carlo simulation combined with confidence intervals (CI) is used to quantify the reliability of the model's predictions. Using a trained GAN model, multiple samplings of the input data are performed by introducing random perturbations (such as input noise or Dropout randomness), generating n sets of predicted values $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N\}$, and the mean $\bar{\hat{y}}$ and standard deviation of the prediction distribution are calculated $\sigma_{\hat{y}}$.

The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to measure the deviation between the model's predicted values and the actual values. MAE calculates the average of the absolute errors between the predicted values and the actual values, and its calculation formula is:

$$\bar{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{6}$$

MAE can intuitively reflect the average level of prediction errors, and the smaller its value, the smaller the average deviation of the model's predictions. RMSE, on the other hand, squares the errors before taking the square root, based on MAE, and its calculation formula is:

$$\sigma_{\hat{y}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

RMSE is more sensitive to larger errors and can reflect the fluctuation degree of prediction errors. Compared with MAE, RMSE amplifies the influence of larger errors, so when there are outliers in the data, the value of RMSE may be greatly affected. Through these two indicators, the prediction accuracy of the model can be evaluated from different perspectives.

The confidence interval coverage is used to measure whether the uncertain interval predicted by the model contains the true value. Suppose for each predicted value, the model provides its corresponding confidence interval [Li, Ui], where Li and Ui are the lower and upper bounds of the confidence interval, respectively. The formula for calculating CI coverage is:

$$\mathrm{CI} = \left[\bar{\hat{y}} - z_{\alpha/2}\cdot\sigma_{\hat{y}}, \bar{\hat{y}} + z_{\alpha/2}\cdot\sigma_{\hat{y}}\right] \tag{8}$$

Here, is the critical value corresponding to the confidence level α of the standard normal distribution (e.g., z0.025≈1.96 for a 95% confidence level). The stability of the predicted values is evaluated based on the width of the confidence interval; a narrower width indicates a more reliable model prediction. Meanwhile, analyzing whether the confidence interval contains the true value helps verify the effectiveness of the model's predictions.

4.3. Comparative Analysis of Core Results

In this study, to evaluate the performance of the proposed PRECISE-Net model in prediction tasks, comparative experiments were conducted with Linear Regression, Long Short-Term Memory (LSTM) networks, and Transformer models. The experimental results are presented using the coefficient of determination (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) on the test set as evaluation metrics, with specific data shown in Table 1 below.

Table.1. Comparison of Model Prediction Accuracy

Model	R^2 (Test Set)	MAE	RMSE
Linear Regression[10]	0.678	2.98	3.96
LSTM	0.729	2.76	3.54
Transformer	0.788	2.43	3.28
PRECISE-Net	0.849	2.08	2.78

For R^2, closer to 1 means better fit. Linear regression has 0.678, LSTM 0.729, Transformer 0.788, and PRECISE-Net 0.849, showing it fits data best. MAE and RMSE measure prediction error; lower values mean higher accuracy. PRECISE-Net has the lowest MAE (2.08) and RMSE (2.78). Overall, PRECISE-Net outperforms other models due to GAN's data generation and feature learning abilities and SA's capture of long-distance dependencies, proving the effectiveness of its adversarial structure and optimization mechanism.

4.4. Analysis of Model Robustness and Hyperparameter Optimization

The robustness of the model and the setting of hyperparameters have a crucial impact on its performance. To enhance the performance of the PRECISE-Net model in this study, some core hyperparameters were optimized and adjusted, and the effects before and after optimization were compared and analyzed through experiments. The relevant hyperparameter information and performance changes are shown in Table 2.

Table.2. Core Hyperparameters

Parameter Name	Default Value	Optimized Value	Performance Improvement
Learning Rate η	0.01	0.001	↑ R^2 +5.2%
Number of Hidden Units	64	256	↑ Reduce Overfitting
Regularization λ	0.01	0.0001	↑ Stabilize Generalization
Time Step T	5	10	↑ Improve Long-Range Dependency Capture

In this study, the optimization of key hyperparameters in the PRECISE-Net model significantly improved its predictive performance and generalization capabilities. The learning rate, initially set at 0.01 and subsequently adjusted to 0.001 post-optimization, enhanced the model's R^2 score by 5.2%. A properly tuned learning rate facilitates precise parameter updates, mitigating risks of overshooting the optimal solution or slow convergence. The reduced learning rate (0.001) allowed for finer parameter adjustments, thereby improving prediction accuracy and model robustness.

The number of hidden units was increased from 64 to 256, which expanded the model's capacity to capture data patterns, thereby reducing underfitting. Despite the inherent risk of overfitting, this adjustment improved generalization performance in the experimental context.

The regularization coefficient, which governs model complexity, was decreased from 0.01 to 0.0001 to stabilize generalization. By penalizing excessive parameter magnitudes in the loss function, this adjustment enabled the model to learn complex features without succumbing to overfitting, thus enhancing its robustness.

The time step parameter, which determines the length of historical information utilized in sequence data, was increased from 5 to 10. This modification improved the model's ability to capture long-range dependencies and enhanced its overall understanding of sequential patterns.

In conclusion, optimizing these core hyperparameters substantially enhanced PRECISE-Net's prediction accuracy, generalization capabilities, and ability to capture long-range dependencies, thereby augmenting its robustness for practical applications.

4.5. Interpretability Analysis (SHAP Explanation)

To explore the contribution of input features to the prediction results in the optimized model, this study introduces the SHAP decomposition method. This method is based on cooperative game theory, calculating the marginal contribution of each feature under different prediction scenarios to quantify its impact on the model's prediction.

Important results are obtained after conducting SHAP analysis on the model. The ranking of the importance of core variables is stable. For example, the SHAP values of features A and B are greater than 0.2, indicating their significant importance. The SHAP dependence plot shows that some variables have a non - linear jumping relationship with the prediction output, and the influence of feature value changes on the prediction results is non - linear. There are obvious linkage and synergistic effects among multiple features. To present the predictive driving forces more intuitively, as illustrated in Figure 4, this study also generated a Local Explainable Model-agnostic Explanations (LIME)-based Force Plot for local sample interpretability. This plot can clearly display the direction and magnitude of each feature's contribution in the prediction of a specific sample, enabling decision-makers to intuitively understand how the model makes predictions based on the input features.

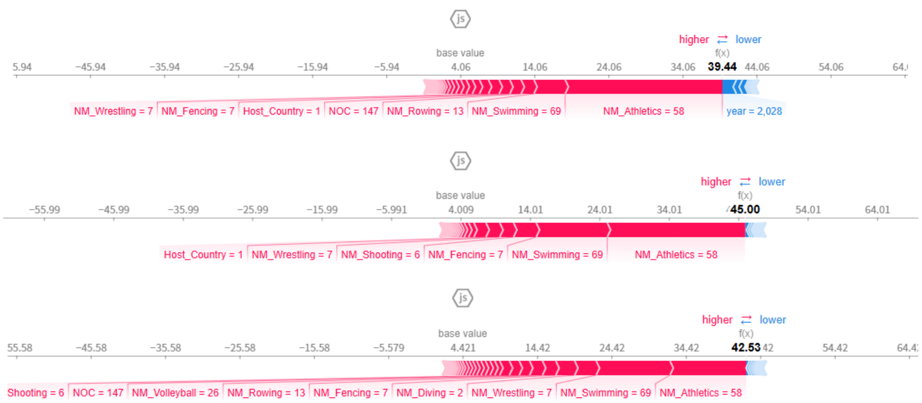

Figure 4. Interpretable Graph of Local Samples in the Model

The analysis results shown in Figure 4 hold multifaceted application value. In terms of feature engineering, it can support feature selection and dimensionality reduction, simplify the model, reduce complexity, and improve generalization ability. In business applications, it provides an interpretation (closed - loop, here it may be better expressed as "an interpretable framework") and decision - making basis for business personnel, helping them understand the business, identify key factors, and formulate reasonable plans. In terms of model interpretability, it achieves a transformation from a "black box" to a "white box", making the decision - making process transparent and enhancing the model's credibility and acceptability.

Therefore, the method proposed in this study not only achieves a leading edge in prediction accuracy but also explains the predictive causality in a visualized manner, revealing the contribution patterns of input variables to the model's output. This characteristic enables the model to have good deployability, transparency, and credibility in actual engineering prediction and auxiliary decision-making systems, providing a powerful tool for solving practical problems.

5. Conclusion

This study proposes an efficient prediction - interpretation framework, the PRECISE - Net model, for the prediction and interpretation of complex problems in fields like engineering. First, in terms of sequence modeling, a sequence modeling method based on Generative Adversarial Networks (GAN) is proposed. Second, to optimize the model structure and enhance generalization ability, this research designs an automatic hyperparameter tuning mechanism integrating Simulated Annealing (SA). Finally, in terms of model interpretability, an interpretability component (SHAP) is introduced to assist in model interpretation. To validate the performance of the proposed framework, this study conducts extensive empirical analyses. By conducting experiments on multiple public datasets and real - world business datasets and comparing the results with mainstream deep models, it is found that the proposed framework outperforms the comparison models on multiple evaluation metrics. These empirical results indicate that the proposed trinity - like general system has practical values such as strong stability, high deployability, and clear interpretation boundaries, providing an effective solution for multi - temporal prediction tasks with high - dimensional structured inputs and complex output targets.

In future research, more advanced optimization strategies will continue to be studied. In terms of model structure, there are plans to integrate Graph Neural Networks (GNN). In terms of temporal modeling capability, it is planned to introduce Transformer variants. In addition, supporting multi - task collaborative prediction and joint optimization is an important direction to address multi - target variable prediction and complex coupling relationships between variables in real - world problems. Meanwhile, to meet the deployment requirements of edge computing scenarios, a lightweight version of the system will be developed, with the model structure optimized, the model compressed, and online learning algorithms studied to achieve dynamic model updates. In the future, the model will continue to be optimized to promote its implementation in more intelligent systems.

References

[1] M. A. Syed, M. Kazerani. A Neural Network-Based Model Predictive Approach for Controlled Electrochemical Green Ammonia Production [C]. IECON 2024 - 50th Annual Conference of the IEEE Industrial Electronics Society, Chicago, IL, USA, 2024: 1-6.

[2] Xinran Liu, Zhenkui Wu. VMD-PCA-LSTM BASED PV POWER PREDICTION MODEL [C]. Proceedings of the 2024 2nd International Conference on Frontiers of Intelligent Manufacturing and Automation (CFIMA '24). Association for Computing Machinery, New York, NY, USA, 2025: 511–516.

[3] Wei Huo, Yu Xiao, Fangyin Lu, Haofu Ji, Tao Liu. Wordle character prediction model based on linear XGBoost [C]. Proceedings of the 2024 2nd International Conference on Communication Networks and Machine Learning (CNML '24). Association for Computing Machinery, New York, NY, USA, 2025: 7–10.

[4] Julio C. Antunez, Johnny D. Salazar, Pedro S. Castañeda. Predictive model based on machine learning for raw material purchasing management in the retail sector [C]. Proceedings of the 2024 International Conference on Advanced Robotics, Automation Engineering and Machine Learning (ARAEML '24). Association for Computing Machinery, New York, NY, USA, 2024: 6–11.

[5] Z. Wang, S. Liu, Y. Pu, S. Chen, T. -Y. Ho, B. Yu. PRO-TIME: Pre-Routing Optimization-Aware Timing Prediction via Multimodal Learning [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2025: 1-1.

[6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative adversarial networks [J]. Commun. ACM, 2020, 63(11): 139–144.

[7] Kirkpatrick S, et al. Optimization by Simulated Annealing [J]. Science, 1983, 220: 671-680.

[8] Scott M. Lundberg, Su-In Lee. A unified approach to interpreting model predictions [C]. Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 2017: 4768–4777.

[9] Hengtong Zhang, Jing Gao, Lu Su. Data Poisoning Attacks Against Outcome Interpretations of Predictive Models [C]. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21). Association for Computing Machinery, New York, NY, USA, 2021: 2165–2173.

[10] M. Huang. Theory and Implementation of linear regression [C]. 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), Chongqing, China, 2020: 210-217.