A Hybrid AI Framework Combining Causal Nonmonotonic Reasoning with VMD-GNN and IGWO-DBN for Decision-Making under Big Data

Hui Li[1*], Chao Xv[2], Ruijuan Liu[2]
[1]Office of Scientific Research, Institute of Higher Education, Jiangxi Polytechnic University, Jiujiang 332007, China
[2]Faculty of Electrical Engineering, Jiangxi Polytechnic University, Jiujiang 332007, China
E-mail: jvtckyc@126.com
[*]Corresponding author

Keywords: Causal relationship, nonmonotonic reasoning, artificial intelligence, big data, DBN

Received: August 30, 2025

With the deep integration of big data and Artificial Intelligence, data-driven predictive models face key challenges such as missing causality and rigid reasoning. Traditional approaches rely on statistical correlations, which struggle to handle uncertainty, conflicting evidence, and sparse data in dynamic environments, resulting in poor interpretability of decision-making. This paper proposes an Artificial Intelligence decision-making model enhanced by causal nonmonotonic reasoning. This model combines Variational Mode Decomposition and Graph Neural Network to process dynamic temporal data and graph structures, and introduces an improved Grey Wolf Optimizer to optimize the parameters of Deep Belief Network, thus constructing a decision framework that integrates dynamic causal discovery, revocable rule inference, and counterfactual evaluation. The experimental results on the MIMIC-III medical dataset and IEEE-CIS financial dataset show that the proposed model achieves an average processing error as low as 0.02, which is superior to the three baseline models Graph Attention Network-Recurrent Neural Network, Particle Swarm Optimization-Ant Colony Optimization, and Genetic Algorithm-Long Short-Term Memory Network. In financial volatility prediction, the model achieved a root mean square error of 0.12, a coefficient of determination of 0.92, and a peak-valley capture rate of 0.89. It also maintained a robust recall rate ranging from 0.8 to 0.98 across confidence intervals of 0 to 0.6, demonstrating the best overall performance. These results indicate that the proposed model shows strong accuracy and adaptability in complex decision-making scenarios. It effectively addresses the problems of causal blindness and rigid reasoning in current Artificial Intelligence systems, offering a new approach to trustworthy Artificial Intelligence and promoting intelligent and efficient development in key fields such as healthcare and finance.

Povzetek: Študija predstavlja model umetne inteligence z vzročnim sklepanjem, ki izboljša natančnost, prilagodljivost in razložljivost odločanja v kompleksnih okoljih.

1 Introduction

In the era of big data, massive data streams are continuously generated by Internet of Things devices, social media, scientific observations, and business systems, forming comprehensive digital mappings of both physical and social systems [1]. Cross-domain data integration technologies are gradually bridging traditional information silos in areas such as healthcare, finance, and industry. This evolving technological ecosystem is driving Artificial Intelligence (AI) from localized pattern recognition toward a new stage of global cognitive intelligence [2]. However, current AI systems rely heavily on data correlations and when facing sparse data, dynamic environments, or adversarial disturbances, these systems exhibit significantly decreased generalization performance [3]. Causal relationships represent stable mechanisms between variables; that is, when an intervention is applied to one variable, it will cause a predictable change in the probability distribution of another variable. This enables AI to go beyond correlation, reveal the mechanisms behind variable interaction, and improve the interpretability and reliability of decisions [4]. Non-monotonic reasoning refers to a computational reasoning method in which new information may overturn original conclusions, thereby providing the ability to dynamically modify inferences, providing AI with the ability to dynamically adapt [5-6]. This study proposes an AI model based on nonmonotonic reasoning and causal relationships. It aims to achieve dynamic recognition, real-time reasoning, and conclusion revision of causal relationships in big data, overcoming the limitations of traditional approaches that depend on static causal structures. Traditional probability models, such as Bayesian networks, rely on prior distributions and fixed structures, making it difficult to maintain stability when data is sparse or dynamically changing. Although existing

explainable AI methods provide transparent decision-making, they lack the ability to learn implicit causal mechanisms. In contrast, the non monotonic reasoning and causal fusion mechanism proposed in this study can dynamically modify assumptions when information is incomplete, effectively alleviating overfitting problems in sparse data. At the same time, the proposed model combines causal discovery and causal graph learning, which can identify and utilize stable causal mechanisms from sparse associations, thus having advantages in dealing with uncertainty and sparse data. The innovation of this study lies in building a dynamic coupling mechanism between defeasible rules and causal graphs. This enables real-time evolution of causal structures and evidence-driven self-correction of conclusions, breaking the rigid constraints of traditional static causal models.

2 Related works

With the development and application of AI, researchers in China and abroad conducted extensive and in-depth studies, including research on causal relationships and nonmonotonic reasoning. Fu et al. proposed a nonmonotonic intermediate-state longitudinal strategy. This process optimized thin-film crystallization and suppressed non-radiative recombination losses, improving the power conversion efficiency in organic solar cells [7]. Sturmberg et al. developed a multi-causal relationship approach to address the multifactorial nature of clinical outcomes. This helped explain the heterogeneity of clinical outcomes more accurately and guide precision interventions [8]. Xie et al. aimed to improve the accuracy of effective connectivity estimation in resting-state functional Magnetic Resonance Imaging (fMRI). They introduced a spectral sampling dynamic causal modeling method, which built a naïve Bayes model in the frequency domain and applied Markov Chain Monte Carlo sampling for parameter estimation. Their method significantly improved the accuracy of estimating directional influences among neural populations and outperformed existing techniques [9]. Komo et al. focused on the limitations of existing systems in nonmonotonic reasoning. They designed a System W non-monotonic reasoning technique, which introduced subset optimization based on conditional sentence influence distribution and a strict partial order representation of worlds. This approach extended the reasoning framework while satisfying additional ideal reasoning properties [10]. Lindström et al. developed a selection function-based semantics for nonmonotonic reasoning to address formalization problems caused by infinite sets of premises. This method used selection functions to filter optimal subsets from possible states and defined inference operations. They also established its axiomatic representation using preference representation theorems [11].

As for the development of methods in the fields of big data and AI, both theoretical and practical applications have become relatively mature. Researchers from various countries have conducted in-depth studies on these topics. Wu et al. proposed a multimodal big data fusion analysis technique. They used AI to integrate multi-scale and multimodal cancer data for deep computational analysis. This approach improved the understanding of disease mechanisms and supported the development of diagnostic and treatment strategies, ultimately enhancing patient outcomes in cancer research [12]. Mainali and his team developed an AI-driven big data analysis method to address the underutilization of massive clinical data in neurocritical care. By applying machine learning, Deep Neural Networks, and efficient data systems, their method processed multi-source information from intensive care units to predict and understand clinical events and outcomes [13]. Lazaroiu's team aimed to optimize the management of cognitive manufacturing systems. They proposed an AI-Internet of Things collaborative technology that integrated AI decision algorithms, sensor networks, and real-time big data analytics. This approach supported sustainable cyber-physical management systems and optimized value creation and production processes in cognitive manufacturing environments [14]. Sohail et al. addressed the challenges of optimizing complex big data. They designed a genetic algorithm optimization method that simulated natural selection to process high-dimensional and stochastic data. This provided an efficient solution for AI and data science applications [15]. Sahu et al. developed an AI-driven big data analysis approach for precision medicine. To solve the challenges in developing individualized treatment plans, their method analyzed multi-omics data to identify patient-specific phenotypes and treatment responses. This promoted a paradigm shift in disease diagnosis, prognosis, and treatment within precision medicine [16]. The summary table of the above related work is shown in Table 1.

Table 1: Summary table of related work

Work	Method	Areas	Key results	Performance gap
[7]	Non monotonic intermediate state longitudinal strategy	Organic solar cells	Power conversion efficiency 19.31%	Lack of a universal causal reasoning framework
[8]	Multiple causal relationships	Clinical outcome analysis	Accurately explain clinical outcomes	Lack of computational models for implementing dynamic causal updates
[9]	Spectrum Sampling Dynamic Causal Model	Resting state fMRI	Accuracy improvement	Cannot undergo non monotonic belief correction with new evidence
[10]	System W Non monotonic Reasoning	Knowledge Representation and Reasoning	Improvement of reasoning characteristics	Lack of ability to learn causal structures from data

[11]	Semantics based on selection function	Non monotonic reasoning	-	Not integrated with specific data-driven AI models
[12]	Multi modal big data fusion AI	Cancer multi omics data	Improved patient outcomes in cancer research	Insufficient interpretability mining of causal mechanisms
[13]	AI driven big data analysis	Intensive Care Unit Data	Clinical event prediction	Limited understanding of causality and adaptability to rules
[14]	AI-Internet of Things collaborative decision-making	Cognitive Manufacturing System	Optimize production process	Lack of explanation on causal discovery and reasoning mechanisms
[15]	Genetic Algorithm	High dimensional random data	Optimize efficiency	Not combined with causal models and non monotonic logic
[16]	AI driven multi omics analysis	Precision Medicine	Develop personalized treatment plan	The interpretability of the decision-making process is low

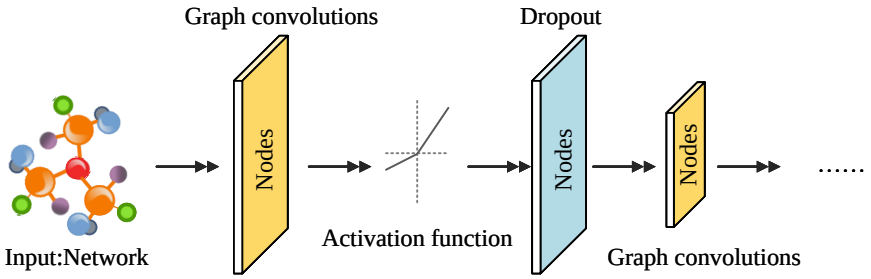


Figure 1: Structure of GNN-based nonmonotonic reasoning design. (Icon source from: https://www.1001freedownloads.com)

In summary, existing research has made progress in data-driven intelligent decision-making. However, most approaches remain limited by static causal assumptions and struggle to adapt to dynamic environments. The flexible rule mechanisms of nonmonotonic reasoning and the traceable intervention capabilities of causal relationships can help address these issues. Therefore, the proposed AI model based on nonmonotonic reasoning and causal relationships is expected to support dynamic evolution of causal structures and real-time self-correction of conclusions, meeting the core demands for robustness, interpretability, and adaptability in high-risk decision-making scenarios. In addition, existing interpretable AI methods often focus on static and post hoc explanations of model decisions, which are frequently disconnected from the causal mechanism and dynamic reasoning process of the model, making it difficult to provide users with truly reliable decision-making basis in complex scenarios. This study aims to construct a framework that can provide dynamic and traceable visual explanations by integrating causal discovery with non-monotonic reasoning, thereby achieving enhanced model interpretability and user trust while maintaining high accuracy.

3 Causal nonmonotonic reasoning and AI Integration under Big Data

3.1 Nonmonotonic reasoning design based on graph neural network algorithm

Traditional deep learning models struggle to capture causal dependencies and evolving rules in dynamic environments, leading to delayed decisions and reduced adaptability. Graph Neural Network (GNN) learns implicit causal representations from high-dimensional graph structures and simultaneously performs evidence-weighted propagation and real-time updates based on defeasible rules [17–18]. Therefore, this study designs a GNN-based nonmonotonic reasoning approach to achieve bidirectional dynamic coupling between causal topology and nonmonotonic rules. The structure is shown in Figure 1.

As shown in Figure 1, the core structure of GNN for nonmonotonic reasoning is built upon dynamic evidence accumulation and belief updating mechanisms based on graph context. It includes three main modules: uncertainty-aware node initialization, multi-layer context-driven evidence aggregation with conflict detection, and conclusion generation with defeasibility output. Each node initializes its belief state based on initial feature vectors and possible confidence scores or probability distributions. The mathematical expression for GNN node state updating is shown in Equation (1).

$$h_v^{(l+1)} = \sigma(\sum_u \in \mathrm{N}(v) \bigcup \{v\} \frac{1}{c_{u,v}} \cdot W^{(l)} h_u^{(l)} + b^{(l)}) \quad (1)$$

In Equation (1), $h_v^{(l)}$ represents the hidden state vector of node v in layer l, $\mathrm{N}(v)$ denotes the set of first-order neighbors of node v, $c_{u,v}$ is the normalization factor, W is the weight matrix, b is the bias vector, and σ is the activation function. The output function of a GNN node is shown in Equation (2).

$$z_v = W^{(L)} h_v^{(L)} + b^{(L)} \quad (2)$$

In Equation (2), z is the final output vector, and L denotes the total number of layers in the network. The graph classification function of GNN is given in Equation (3).

$$h_G = \rho(\{h_v^{(L)} | v \in v'\}) \quad (3)$$

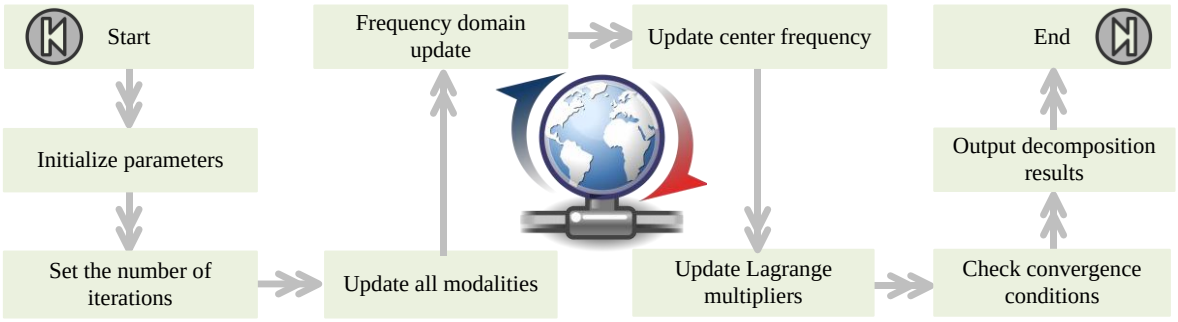


Figure 2: Structure of VMD applied to nonmonotonic reasoning design. (Icon source from: https://www.1001freedownloads.com)

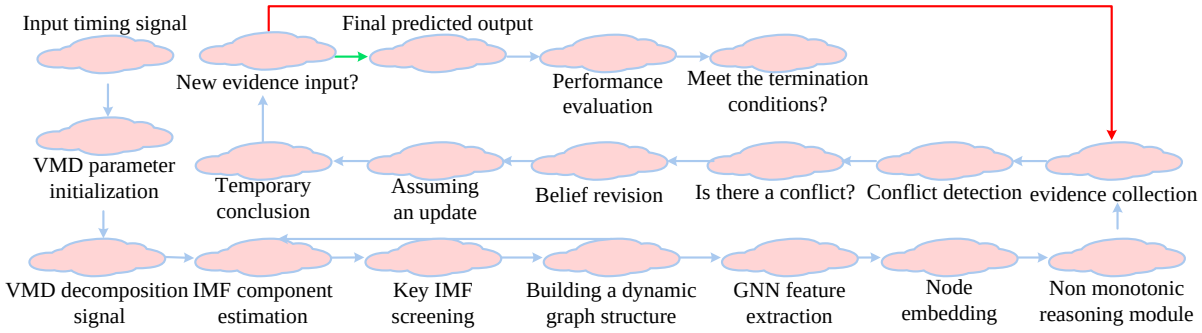


Figure 3: Structure of VMD-GNN

In Equation (3), h_G represents the global representation vector of the graph, h is the hidden state of each node, v' is the set of all nodes, and ρ is the aggregation function. GNN is prone to gradient vanishing and feature confusion when processing long-term dependencies and high-frequency noise data. The introduction of Variational Mode Decomposition (VMD) directly addresses the limitations of GNN in processing raw temporal data. VMD decomposes complex signals into a series of smooth, narrowband intrinsic mode functions through adaptive band partitioning, thereby transforming the frequency domain features of the temporal data into multi-scale spatial structural features that can be processed by GNN [19-20]. The structure of VMD applied to nonmonotonic reasoning is shown in Figure 2.

As shown in Figure 2, the VMD process starts with parameter initialization. The system sets the basic parameters of the reasoning model and initializes all dynamic hypotheses with their initial confidence levels and constraint balancing factors. The iteration counter is also set to zero. In the main iteration loop, the counter increases by one in each round. Then, all modal components are updated in parallel in the frequency domain. Each mode is reconstructed using specific frequency-domain operations defined by the evidence fusion rules. The center frequency of each mode is then recalculated and updated as the spectral energy centroid. Finally, the constraint balancing factor is updated to maintain internal consistency and reinforce decomposition constraints. After each round of updates, the system checks the convergence condition. If the change in modal components is below the threshold, the loop ends and the final decomposition result is output. This iterative process of hypothesis updating, confidence re-evaluation, and constraint balancing simulates the mechanism of new evidence in nonmonotonic reasoning, allowing conclusions to be dynamically revised until stabilization. The mathematical expression of the VMD variational problem is given in Equation (4).

$$\min_{\{u_k\},\{w_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\delta(t) + \frac{j}{\pi t} * u_k(t) \right] e^{-jwkt} \right\|_2^2 \right\} \quad (4)$$

In Equation (4), u_k denotes the decomposed signal component, w_k is the center frequency, K is the predefined number of modes, $\delta(t)$ is the Dirac function, ∂_t is the partial derivative with respect to time t, and j is the imaginary unit. The core objective of equation (4) is to decompose a complex signal into multiple simple sub signals, effectively separating the different causal components at various scales that are mixed together. The constraint condition of VMD is expressed in Equation (5).

$$\sum_{k=1}^{K} u_k(t) = f(t) \quad (5)$$

In Equation (5), $f(t)$ represents the constraint condition. Optimizing GNN using VMD forms the VMD-GNN model, whose structure is shown in Figure 3.

As shown in Figure 3, the VMD-GNN process begins with inputting a time series signal. The system initializes VMD parameters and decomposes the signal iteratively to

obtain intrinsic mode functions. After evaluating and selecting key components, it builds a dynamic graph structure. GNN then performs feature extraction and node embedding on the graph. Next, it enters the core nonmonotonic reasoning module, which includes evidence collection, conflict detection, and belief revision mechanisms. This module dynamically handles conflicting evidence and iteratively updates hypotheses, generating final prediction outputs through verification loops. A performance evaluation module determines whether stopping conditions are met. If not, the system triggers a parameter adjustment mechanism to re-optimize the VMD decomposition process, forming a closed feedback loop. The mathematical expression for updating the modal function in VMD is shown in Equation (6).

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum_{i\neq k}\hat{u}_i(\omega)+\frac{\hat{\lambda}(\omega)}{2}}{1+2\alpha(\omega-\omega_k)^2} \tag{6}$$

In Equation (6), ω denotes the Fourier transform, and n represents the number of iterations. The equation for updating the center frequency in VMD is shown in Equation (7).

$$\omega_k^{n+1}=\frac{\int_0^\infty \omega\left|\hat{u}_k(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_k(\omega)\right|^2 d\omega} \tag{7}$$

In Equation (7), ω_k represents the new center frequency of the mode.

3.2 Construction of nonmonotonic reasoning model with integrated causal relationships

Although VMD-GNN effectively handles defeasible reasoning in complex dynamic rules and graph structure evolution, it still has limitations in extracting deep causal features and modeling cross-scale dependencies. DBN offers hierarchical feature abstraction and probabilistic generative capabilities, which capture latent causal factors and high-order nonlinear associations in data. Therefore, this study introduces DBN based on the nonmonotonic reasoning method of VMD-GNN. A multi-level joint inference framework combining features, rules, and causality is constructed. The energy function of the restricted Boltzmann machine in DBN is shown in Equation (8).

$$E(b,g)=-\sum_i a_i b_i-\sum_j m_j g_j-\sum_{ij} bW_{ij}g_j \tag{8}$$

In Equation (8), b represents the vector state of the visible layer, j denotes the value of the units, and g represents the vector state of the hidden layer. Equation (8) measures the degree of harmony between a certain state configuration between the visible and hidden layers of a neural network, and patterns that conform to causal laws have a higher probability of appearing in the system. The DBN structure integrated with causal relationships is shown in Figure 4.

As shown in Figure 4, the DBN structure consists of an input unit, a causal feature extraction unit, and a causality-aware output unit. Data flow from the input unit to the causal feature extraction unit and finally to the output unit. The input unit includes three nodes, which are fully connected to the first feature extraction group with four nodes. These nodes output to the causal mechanism modeling layer to identify potential causal relationships among variables. The extracted causal features are then fully connected to the second feature extraction group with three nodes and finally linked to the output node. This network structure achieves hierarchical extraction and combination of observed and causal features. It first performs unsupervised layer-wise learning of statistical associations and latent causal patterns, then fine-tunes network parameters through supervised learning with causal constraints. Fully connected layers between units ensure sufficient information flow and integration. The joint concept distribution in DBN is shown in Equation (9).

$$P(b,g)=\frac{1}{Z}e^{-E(b,g)} \tag{9}$$

In Equation (9), Z denotes the partition function. The conditional probability distribution of DBN is calculated as shown in Equation (10).

$$P(g_j=1|b)=\sigma(b_j+\sum_i b_i W_{ij}) \tag{10}$$

In Equation (10), $P(g_j=1|b)$ represents the activation of the hidden layer. The activation probability of the visible layer in DBN is expressed in Equation (11).

$$P(b_j=1|b)=\sigma(a_j+\sum_i g_i W_{ij}) \tag{11}$$

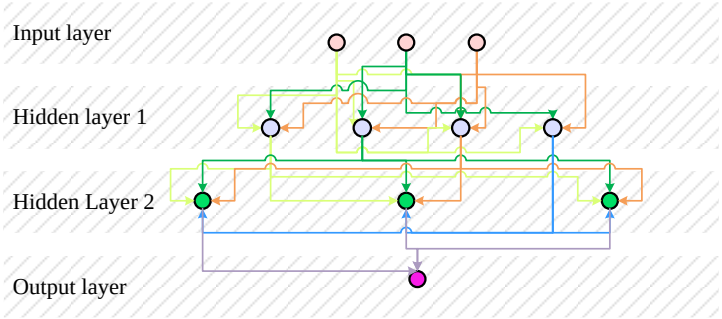


Figure 4: Structure of DBN integrated with causal relationships

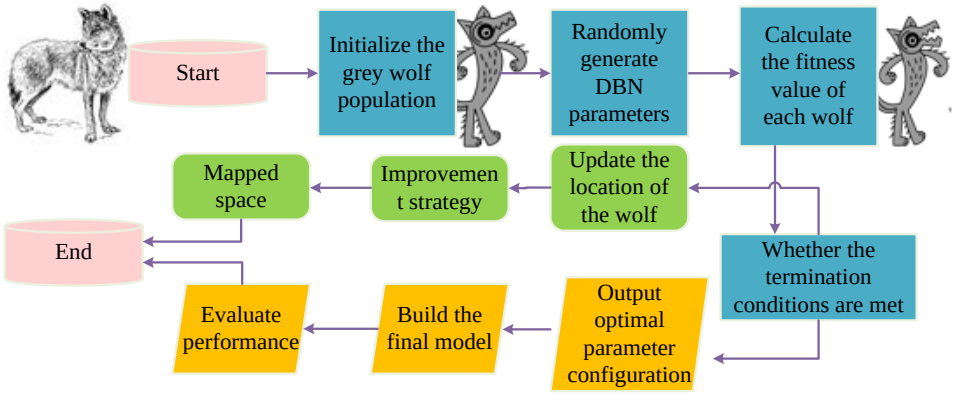


Figure 5: Flowchart of IGWO-DBN. (Icon source from: https://www.1001freedownloads.com)

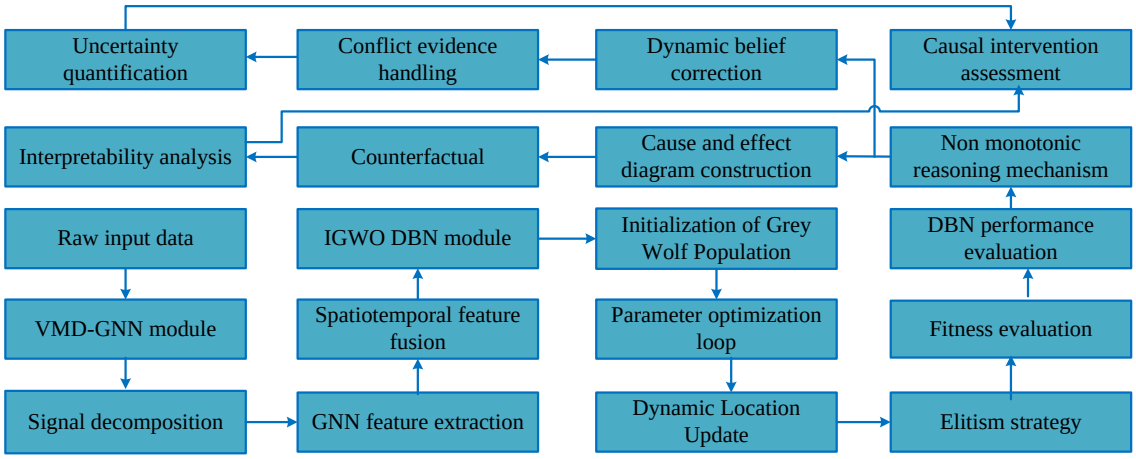


Figure 6: The AI model based on nonmonotonic reasoning and causal relationships

In Equation (11), $P(b_j = 1|b)$ denotes the activation of the visible layer. DBN tends to fall into local optima and has low training efficiency. IGWO, through swarm intelligence optimization, effectively escapes local optima and accelerates convergence. By optimizing DBN with IGWO, the IGWO-DBN is obtained. Its flow is shown in Figure 5.

As shown in Figure 5, the IGWO-DBN process starts at the beginning node. First, the grey wolf population is initialized, and DBN parameters are randomly generated. Then, the fitness value of each wolf is calculated. If the termination condition is not met, the population positions are updated and enhanced strategies are applied. New positions are mapped to the DBN parameter space, and the fitness values are recalculated. Once the termination condition is satisfied, the optimal parameter configuration is output. The final DBN model is constructed using these parameters and its performance is evaluated on the test set, leading to the end node. The encircling prey equation of IGWO is shown in Equation (12).

$$D = \left| C \cdot X_p(n) - X(n) \right| \tag{12}$$

In Equation (12), X_p is the position of the prey, D denotes the position of the grey wolf, and C is the coefficient vector. The fitness function of IGWO is shown in Equation (13).

$$\Re = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)) \tag{13}$$

In Equation (13), y_i represents the true labels of the samples, and N denotes the number of training samples. VMD-GNN and IGWO-DBN are combined to form an AI model based on nonmonotonic reasoning and causal relationships. Its structure is shown in Figure 6.

In Figure 6, the combination of GNN based modules and DBN based modules achieves functional complementarity and collaboration. The core contribution of the GNN module lies in its topological awareness and dynamic inference capabilities, which can directly aggregate neighborhood information and update real-time beliefs in graph structures. The core contribution of the DBN module lies in its deep causal representation ability, which can abstract stable high-order causal factors and generation mechanisms from data through hierarchical nonlinear transformations. Each component forms an organic whole through close collaboration. VMD, as a front-end processor, ensures the data quality of the input to the GNN. GNN is responsible for handling dynamically evolving graph contexts. IGWO serves as an optimizer to ensure efficient convergence of DBN. DBN is responsible for mining the stability mechanisms behind the data.

4 Performance of the AI Model based on nonmonotonic reasoning and causal relationships

4.1 Effectiveness verification of VMD-GNN

To verify the superiority of VMD-GNN, it was compared with Graph Convolutional Network-Long Short-Term Memory (GCN-LSTM), Bayesian Defeasible Reasoning Network (BDRN), and Neurosymbolic Dynamic System (NSDS). GCN-LSTM represents an advanced deep learning method for modeling spatiotemporal dynamics, BDRN is a typical representative of Bayesian probability models for handling uncertainty, and NSDS reflects the latest progress in neural symbol ensemble. By comparing with models of these three different paradigms, the effectiveness of VMD-GNN in basic feature extraction and dynamic rule learning can be comprehensively tested. The experimental system used Ubuntu 20.04 LTS, with Linux Kernel 5.15, PyTorch 1.12 + PyG 2.2 as the deep learning framework, AdamW as the optimizer, and Python 3.9 as the programming language. The hardware included an NVIDIA A100 80GB PCIe GPU and 512GB DDR4 RAM. To ensure the realism and reliability of the experiments, the RSS3 Web3 dataset and the MNIST dataset were used for training and testing. The RSS3 Web3 and MNIST datasets provide complex graph structure relationships and standard image features, respectively, which can verify the universal ability of the model in basic feature extraction and dynamic rule learning. The RSS3 Web3 dataset is divided into training, testing, and validation sets in a ratio of 7:2:1. The MNIST dataset uses 60000 images as the training set and 10000 images as the testing set, with pixel values normalized to [0,1]. All experiments were independently repeated 5 times, initialized with fixed random seeds (42, 123, 456, 789, 101, 112) to control randomness. The key hyperparameters of the model were determined through grid search, with a VMD mode number of 5, a penalty factor of 2000, a GNN hidden layer dimension of 128, and a learning rate of 0.001. The comparison of update steps and loss values is shown in Figure 7.

In Figure 7(a), the initial loss of the proposed algorithm on the RSS3 Web3 dataset was 0.31. After 100 iterations, it quickly dropped to 0.12 and eventually stabilized between 0.02 and 0.05, significantly outperforming the baseline models. Figure 7(b) shows that on the MNIST dataset, the initial loss of the proposed algorithm was 0.25. It decreased to 0.15 after 100 iterations and finally converged to 0.08, again far better than the baselines. In summary, the proposed algorithm completed major optimization within the first 30% of training epochs on both datasets, achieving 2–3 times higher training efficiency than the comparison models. The experiments demonstrated that the model effectively addressed core challenges of path planning in dynamic conditions. To further validate the effectiveness of VMD-GNN, its accuracy and loss rate were compared, as shown in Figure 8.

As shown in Figure 8(a), the accuracy of the proposed algorithm increased rapidly to 75% before 60 iterations. Between 60 and 90 iterations, the improvement slowed but continued. After 150 iterations, the accuracy stabilized above 90% and approached 100%. Although GCN-LSTM performed better than the proposed algorithm before 60 iterations, it showed significant fluctuations afterward, leading to inferior overall accuracy. Figure 8(b) shows that the loss rate of the proposed algorithm dropped quickly and converged to 0.9. The total loss remained under 0.10. After 200 iterations, the curve continued to decline steadily, nearing 1. In summary, the proposed algorithm achieved an average test accuracy of 93% and an average loss of 0.11, offering better convergence precision and stability than the baselines. To further assess VMD-GNN's classification performance, it was compared with other models in terms of Area Under Curve (AUC) and F1 score. The results are shown in Figure 9.

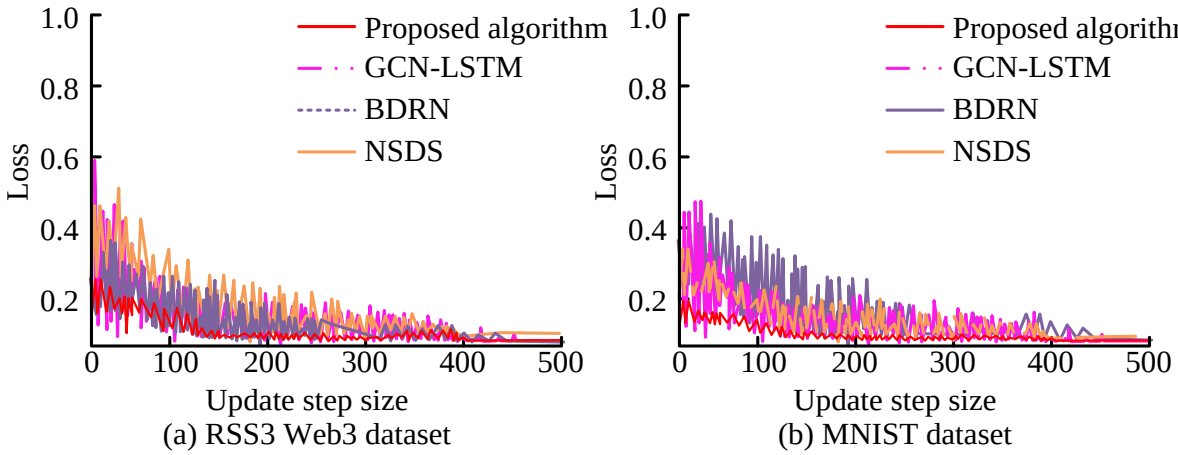


Figure 7: Comparison of update steps and loss values for four models

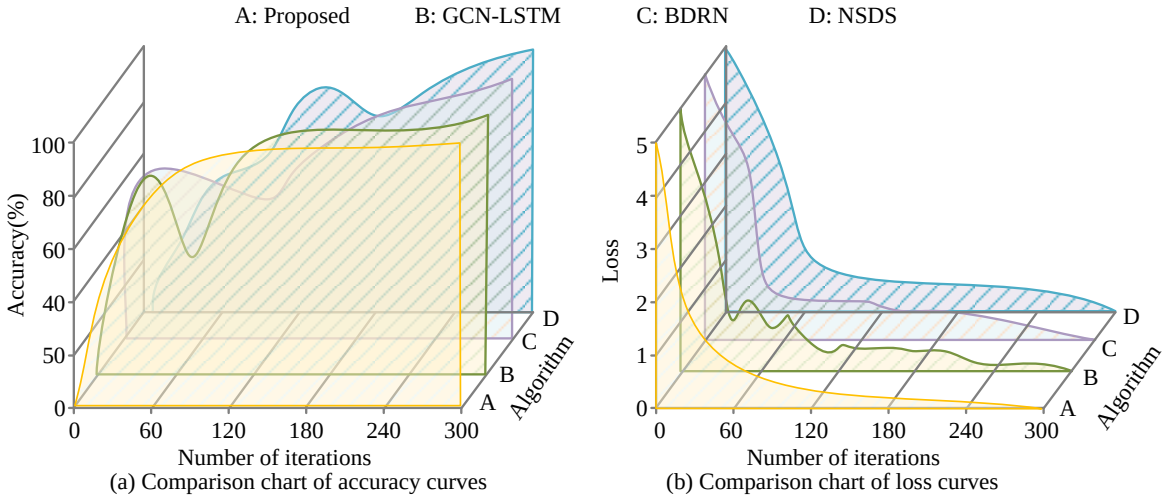


Figure 8: Comparison of accuracy and loss rate for four models

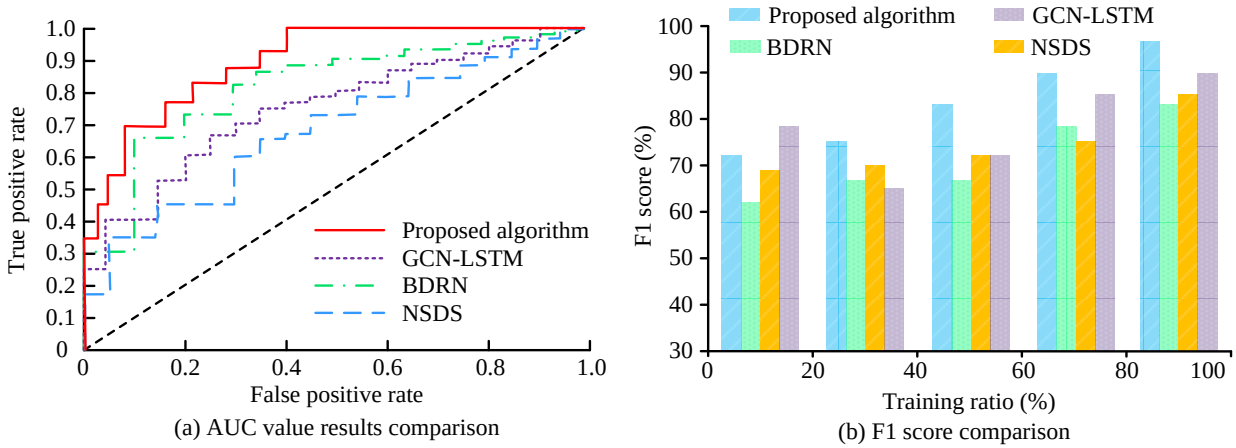


Figure 9: ROC curve and F1 score comparison for four models.

Figure 9(a) shows that the Receiver Operating Characteristic (ROC) curve of the proposed algorithm was closest to the top-left corner. Its AUC reached 0.92, significantly higher than the comparison models' 0.81, 0.82, and 0.78, indicating better classification ability and prediction reliability. Figure 9(b) shows that the F1 scores of the proposed algorithm remained above 80%, much higher than those of the baseline models, further confirming its classification performance. In summary, the proposed algorithm outperformed others in classification prediction and recall, showing stronger robustness.

4.2 Application evaluation of the integrated causal nonmonotonic reasoning and AI model

The study aimed to verify the following core hypothesis: whether the performance and robustness of the integrated model incorporating non-monotonic reasoning and causal relationships in causal effect estimation and prediction tasks in dynamic and high noise environments are significantly better than existing representative baseline models. The manipulated independent variable in the experiment is the model architecture, and the observed dependent variables include prediction accuracy, estimation robustness, regression indicators, and computational efficiency. Distributed training was conducted using PyTorch Lightning + DeepSpeed. The development environment was NVIDIA NGC 22.08 container, with dual Intel Xeon Platinum 8380 CPUs. The datasets used were MIMIC-III and IEEE-CIS. The MIMIC-III dataset consists of rich longitudinal electronic health records of patients, including sparse, incomplete, and confounding clinical data. This provides an ideal scenario for validating the causal discovery and non-monotonic reasoning ability of the model in the face of evidence conflicts. The IEEE-CIS financial transaction dataset exhibits significant concept drift, high class imbalance, and adversarial noise. Its essence lies in identifying causal relationships between fraudulent behavior and transaction risks, making it highly suitable for testing models' ability to maintain causal inference stability in dynamic environments. On the basis of

dividing by patient ID, the MIMIC-III dataset normalizes all continuous variables with Z-score, encodes categorical variables with one hot encoding, and processes missing values. The IEEE-CIS dataset is divided into training, testing, and set validation in a 7:2:1 ratio. The DBN structure was set to [256, 128, 64], the IGWO population size to 30, and the maximum iterations to 200. First, to verify the effectiveness of each core component in the improved integrated model, ablation experiments were conducted on the IEEE-CIS dataset. VMD, GNN, IGWO, and DBN were removed separately and compared with the complete model. The results are shown in Table 2.

From Table 2, it can be seen that removing VMD, GNN, IGWO, and DBN all led to a significant decrease in the model's metric results ($p<0.05$). The performance loss caused by removing DBN is the most severe, with Macro-F1 decreasing to 0.81 and RMSE increasing to 0.24 ($p<0.01$). This highlights the core position of DBN in deep causal feature abstraction and high-order nonlinear relationship modeling. The results indicated that each component of the proposed model was of great significance for achieving high-precision causal inference in complex dynamic environments. Using Graph Attention Network-Recurrent Neural Network (GAT-RNN), Particle Swarm Optimization-Ant Colony Optimization (PSO-ACO), and Genetic Algorithm-Long Short-Term Memory Network (GA-LSTM) as baselines. GAT-RNN represents an advanced deep learning method for modeling spatiotemporal dynamics, PSO-ACO is a typical combination of metaheuristic optimization algorithms, and GA-LSTM represents a class of methods that integrate genetic algorithms and deep learning to optimize temporal models. Compared with such strong baselines, it strongly demonstrated the comprehensive performance improvement of the proposed model after integrating causal reasoning and non-monotonic logic. The Average Treatment Effect (ATE) comparison of the four models is shown in Figure 10.

Figure 10(a) shows that on the MIMIC-III dataset, the ATE of the proposed model remained optimal throughout, consistently at 0.02 with zero fluctuation as the sample size increased from 5,000 to 25,000. GAT-RNN's ATE rose from 0.05 to 0.15 and stayed at that level. PSO-ACO fluctuated from 0.22 to 0.31 and then dropped to 0.19, while GA-LSTM exhibited large fluctuations throughout. Figure 10(b) shows that on the IEEE-CIS dataset, the proposed model again demonstrated superior performance, maintaining an ATE of 0.02 with no drift as the sample size grew. GAT-RNN's ATE dropped from 0.18 to 0.07, PSO-ACO rose slowly from 0.12 to 0.14, while others showed instability. The robust ATE of the proposed model stemmed from the synergistic effect of its causal and non-monotonic components. When the distribution of patient subpopulation characteristics changes, non-monotonic reasoning allows the model to dynamically adjust the rule set used for effect estimation, while causal graphs ensure that this dynamic adjustment is based on the correct causal structure, avoiding biased estimation. In contrast, the baseline models lacked this adaptive causal understanding ability, resulting in significant fluctuations in ATE estimation. To further validate the effectiveness of the proposed model, its recall rate was compared with other models, as shown in Figure 11.

Table 2: Results of ablation experiment

Models	Macro-F1	AUC	RMSE
-VMD	0.86±0.03**	0.87±0.02**	0.18±0.04**
-GNN	0.84±0.04**	0.86±0.02**	0.20±0.05**
-IGWO	0.87±0.03*	0.89±0.01*	0.15±0.03*
-DBN	0.81±0.04**	0.85±0.02**	0.24±0.06**
Complete model	0.91±0.01	0.92±0.01	0.12±0.02

Note: * indicates significant differences between the model and the complete model, $p<0.05$, ** indicates a highly significant difference between the model and the complete model, $p<0.01$.

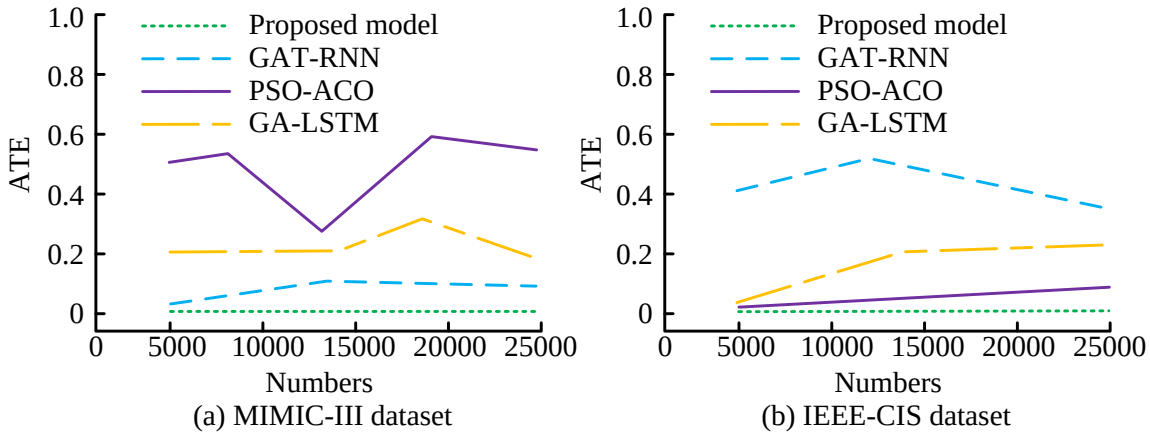
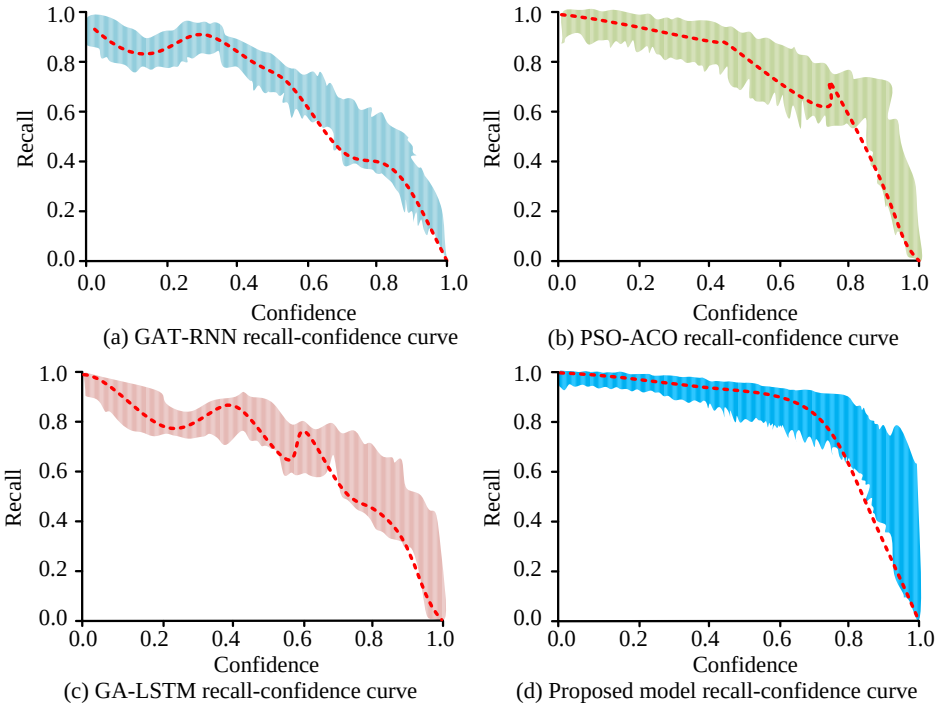


Figure 10: ATE comparison for four models



(a) GAT-RNN recall-confidence curve (b) PSO-ACO recall-confidence curve

(c) GA-LSTM recall-confidence curve (d) Proposed model recall-confidence curve

Figure 11: Recall comparison for four models

As shown in Figure 11, the GAT-RNN model showed a three-phase decay pattern: it dropped from 1.0 to 0.82 in the 0–0.4 confidence interval, declined to 0.6 in the 0.4–0.6 range, and then dropped to zero. PSO-ACO maintained a high recall above 0.95 in the 0–0.4 range, dropped sharply to 0.7 between 0.4–0.6, and then dropped to zero. GA-LSTM performed the worst, falling to 0.8 in the 0–0.2 range, and experienced a cliff drop from 0.6 to 0.2 between 0.6–0.8. In contrast, the proposed model showed unique robustness. Its recall rate slowly decreased from 0.98 to 0.8 within the low confidence interval range of 0-0.6, and remained stable at 0.5 within the high confidence interval range of 0.6-0.8 without further decay. This reflected a reliable and conservative decision-making strategy of the model. In the low confidence interval, the model maintained a high recall rate to capture potential targets widely. When the confidence level reached 0.6 or above, the model only provided high confidence predictions for instances supported by strong causal evidence, actively filtering out boundary cases with high uncertainty, ensuring a significant improvement in prediction accuracy within the high confidence interval. To further validate the effectiveness of the proposed model, its prediction performance in financial volatility forecasting was compared with other models, as shown in Table 3.

As shown in Table 3, the proposed model achieved a significantly lower Root Mean Squared Error (RMSE) of 0.12 compared to the baselines, and its Mean Absolute Error (MAE) of 0.08 represented a 38.5% reduction from the next best model. The Coefficient of Determination (R^2) reached 0.92, and directional prediction accuracy was also 0.92. For key event detection, the peak-valley detection rate was 0.89, exceeding the baselines by 11–18 percentage points. The F1 score for extreme events reached 0.91, leading by 9–15 points. The model also demonstrated outstanding real-time performance, with volatility lag of only 8.2 ms. This outstanding predictive performance was directly attributed to its causal and non-monotonic reasoning core components. The causal diagram in the proposed model could identify and rely on relatively stable causal mechanisms between variables, thereby capturing the essential driving forces in changes. At the same time, non-monotonic inference components allowed the model to dynamically undo old patterns that have expired and activate inference rules that are adapted to the new environment, thereby maintaining high accuracy in high noise and concept drift environments. In order to further identify the performance advantages of the proposed model, it was compared with three representative methods, namely causal forest model, proximal policy optimization (PPO), and XGBoost, on the IEEE-CIS dataset. Macro F1 score, AUC, and peak detection delay are used as evaluation indicators. The comparison results of the indicators of the four methods are shown in Table 4.

Table 3: Prediction performance comparison in financial volatility forecasting

Evaluation indicators	Proposed model	GAT-RNN	PSO-ACO	GA-LSTM
RMSE	0.12	0.18	0.21	0.19
MAE	0.08	0.13	0.15	0.14
R2	0.92	0.86	0.81	0.84
Direction Accuracy	0.92	0.85	0.82	0.86
Peak valley capture rate	0.89	0.76	0.71	0.78
Extreme event F1 score	0.91	0.82	0.76	0.79
Volatility lag (ms)	8.2	35.7	42.3	28.5
Word reasoning consumes energy (J)	0.42	1.37	2.05	0.98

Table 4: Comparison results of indicators for four methods

Methods	Macro-F1	AUC	Peak Latency/ms
Causal Forest	0.84	0.88	22.53
PPO	0.81	0.86	105.71
XGBoost	0.87	0.89	15.34
Proposed model	0.90	0.92	8.19

From Table 4, it can be seen that compared to the causal forest model, PPO, and XGBoos, the proposed model performs better in macro F1 score, AUC, and peak detection delay indicators, with values of 0.90, 0.92, and 8.19ms, respectively. The results indicated that the proposed model successfully combined the stability of causal reasoning, the adaptability of non monotonic logic, and the high expressiveness of deep learning, achieving comprehensive performance improvement in complex scenarios.

5 Discussion

To address the challenges of causality absence and rigid reasoning faced by traditional AI models in dynamic big data environments, this study proposed an AI model based on nonmonotonic reasoning mechanisms and causal relationships. Experimental results showed that the proposed model achieved the best performance in ATE. In the recall rate test, the model's score dropped from 0.98 to 0.8 within the 0–0.6 confidence interval, with a decay rate 57% lower than PSO-ACO and 82% slower than GA-LSTM. In financial forecasting tasks, the model reached an RMSE of 0.12, MAE of 0.08, and R^2 of 0.92. It also led in key metrics such as peak-trough capture rate (0.89) and extreme event F1 score (0.91). The model demonstrated clear advantages in computational efficiency, with inference latency of 8.2 ms and energy consumption of 0.42 J—only 20.5% of PSO-ACO. The proposed model exhibits superior performance in both ATE stability and macro F1 score. Compared to Xie et al.'s [9] who focused on static causal estimation methods and Mainali et al. [13] who focused on clinical data association prediction models, the advantage of the proposed model lies in its use of dynamic causal graphs to capture stable causal mechanisms between variables, while non-monotonic inference components allow the system to dynamically revoke failed rules and activate new rules when faced with new evidence or concept drift, thereby maintaining high accuracy and strong robustness in changing environments. Compared to Boulkroune et al.'s [21] adaptive fuzzy control, the proposed model achieves adaptive recognition of unmodeled dynamics through data-driven causal discovery and revocable logic. Compared to Zouari et al.'s [22] robust neural adaptive control, this framework provides a clear decision path through causal diagrams and non-monotonic reasoning while maintaining learning ability. Although Rigatos et al. [23] made progress in nonlinear optimal control of gas compressors, their method did not involve the discovery and updating of system causal mechanisms. Therefore, the proposed model can achieve intelligent decision-making with greater interpretability and environmental robustness than traditional control methods, providing a new approach for industrial systems with strong nonlinearity.

The proposed model framework has domain generality, as it utilizes mechanisms such as VMD and causal stability to handle high-dimensional noise, as well as non-monotonic reasoning to handle data sparsity and conflicts. It is more effective than traditional static models in extracting reliable causal conclusions from noisy high-dimensional data in the dynamic healthcare and social network fields. However, this powerful performance was accompanied by an increase in computational complexity and hyperparameter adjustment requirements, such as the need to optimize the modal number K of VMD and the network structure of DBN. The current main limitation lies in the dependency on domain knowledge for causal graph initialization, which to some extent restricted its rapid deployment in new scenarios. Therefore, future research will focus on developing automated hyperparameter search strategies, combined with incremental learning paradigms such as federated learning, to lower deployment thresholds while maintaining performance.

6 Conclusion

In summary, the model exhibited strong adaptability and interpretability, and met core demands for reliability and transparency in high-risk scenarios. Despite its outstanding performance in dynamic causal discovery and rule conflict resolution, the study still faced limitations in terms of high computational complexity and strong dependence on hyperparameters. Future research will explore improving real-time capability through distributed incremental learning algorithms and will aim to further

enhance the model's prediction generalization and effectiveness.

References

[1] Mehdi Gheisari, Hooman Hamidpour, Yang Liu, Peyman Saedi, Arif Raza. Ahmad Jalili, Hamidreza Rokhsati, and Rashid Amin. Data mining techniques for web mining: A survey. Artificial Intelligence and Applications, 1(1):3-10, 2023. https://doi.org/10.47852/bonviewAIA2202290

[2] Kun-Hsing Yu, Elizabeth Healey, Tze-Yun Leong, Isaac S Kohane, and Arjun K Manrai. Medical artificial intelligence and human values. New England Journal of Medicine, 390(20):1895-1904, 2024. https://doi.org/10.1056/NEJMra2214183

[3] Joseph Uscinski, Adam M. Enders, Casey Klofstad, and Justin Stoler. Cause and effect: On the antecedents and consequences of conspiracy theory beliefs. Current Opinion in Psychology, 47(2):101364, 2022. https://doi.org/10.1016/j.copsyc.2022.101364

[4] Augustine Kumah, Chinwe N Nwogu, Abdul-Razak Issah, Emmanuel Obot, Deborah T Kanamitie, Jerry S Sifa, and Lawrencia A Aidoo. Cause-and-effect (fishbone) diagram: A tool for generating and organizing quality improvement ideas. Global Journal on Quality and Safety in Healthcare, 7(2):85-87, 2024. https://doi.org/10.36401/JQSH-23-42

[5] Matías Osta-Vélez, and Peter Gärdenfors. Nonmonotonic reasoning, expectations orderings, and conceptual spaces. Journal of Logic, Language and Information, 8(3):44-65, 2022. https://doi.org/10.1007/s10849-021-09347-6

[6] Pei-Chi Lo, and Ee-Peng Lim. Non-monotonic generation of knowledge paths for context understanding. ACM Transactions on Management Information Systems, 15(1):1-28, 2024. https://doi.org/10.1145/3627994

[7] Jiehao Fu, Patrick W. K. Fong, Heng Liu, Chieh-Szu Huang, Xinhui Lu, Shirong Lu, Maged Abdelsamie, Tim Kodalle, Carolin M. Sutter-Fella, Yang Yang, and Gang Li. 19.31% binary organic solar cell and low non-radiative recombination enabled by non-monotonic intermediate state transition. Nature Communications, 14(1):1760-1771, 2023. https://doi.org/10.1038/s41467-023-37526-5

[8] Joachim P Sturmberg, and James A Marcum. From cause and effect to causes and effects. Journal of Evaluation in Clinical Practice, 30(2):296-308, 2024. https://doi.org/10.1111/jep.13814

[9] Yuhai Xie, Puming Zhang, and Jun Zhao. A spectral sampling algorithm in dynamic causal modelling for resting-state fMRI. Human Brain Mapping, 44(8):2981-2992, 2023. https://doi.org/10.1002/hbm.26256

[10] Christian Komo, and Christoph Beierle. Nonmonotonic reasoning from conditional knowledge bases with system W. Annals of Mathematics and Artificial Intelligence, 90(1):107-144, 2022. https://doi.org/10.1007/s10472-021-09777-9

[11] Sten Lindström. A semantic approach to nonmonotonic reasoning: inference operations and choice. Theoria, 88(3):494-528, 2022. https://doi.org/10.1111/theo.12405

[12] Xifeng Wu, Wenyuan Li, and Huakang Tu. Big data and artificial intelligence in cancer research. Trends in Cancer, 10(2):147-160, 2024. https://doi.org/10.1016/j.trecan

[13] Shraddha Mainali, and Soojin Park. Artificial intelligence and big data science in neurocritical care. Critical Care Clinics, 39(1):235-242, 2022. https://doi.org/10.1016/j.ccc.2022.07.008

[14] George Lăzăroiu, Armenia Androniceanu, Iulia GRECU, Gheorghe Grecu, and Octav Neguriță. Artificial intelligence-based decision-making algorithms, Internet of Things sensing networks, and sustainable cyber-physical management systems in big data-driven cognitive manufacturing. Oeconomia Copernicana, 13(4):1047-1080, 2022. https://doi.org/10.24136/oc.2022.030

[15] Ayesha Sohail. Genetic algorithms in the fields of artificial intelligence and data sciences. Annals of Data Science, 10(4):1007-1018, 2023. https://doi.org/10.1007/s40745-021-00354-9

[16] Mehar Sahu, Rohan Gupta, Rashmi K Ambasta, and Pravir Kumar. Artificial intelligence and machine learning in precision medicine: A paradigm shift in big data analysis. Progress in Molecular Biology and Translational Science, 190(1):57-100, 2022. https://doi.org/10.1016/bs.pmbts.2022.03.002

[17] Tao Chen, Richang Hong, Yanrong Guo, Shijie Hao, and Bin Hu. MS²-GNN: Exploring GNN-based multimodal fusion network for depression detection. IEEE Transactions on Cybernetics, 53(12):7749-7759, 2022. https://doi.org/10.1109/TCYB.2022.3197127

[18] Shuke Zhang, Yanzhao Jin, Tianmeng Liu, Qi Wang, Zhaohui Zhang, Shuliang Zhao, and Bo Shan. SS-GNN: A simple-structured graph neural network for affinity prediction. ACS Omega, 8(25):22496-22507, 2023. https://doi.org/10.1021/acsomega.3c00085

[19] Mariano Spivak, John E. Stone, João Ribeiro, Jan Saam, Peter L. Freddolino, Rafael C. Bernardi, and Emad Tajkhorshid. VMD as a platform for interactive small molecule preparation and visualization in quantum and classical simulations. Journal of chemical information and modeling, 63(15):4664-4678, 2023. https://doi.org/10.1021/acs.jcim.3c00658

[20] Jérôme Hénin, Laura J. S. Lopes, and Giacomo Fiorin. Human learning for molecular simulations: The collective variables dashboard in VMD. Journal of Chemical Theory and Computation, 18(3):1945-1956, 2022. https://doi.org/10.1021/acs.jctc.1c01081

[21] Abdesselem Boulkroune, Farouk Zouari, and Amina Boubellouta. Adaptive fuzzy control for practical fixed-time synchronization of fractional-order chaotic systems. Journal of Vibration and Control,

10775463251320258, 2025. https://doi.org/10.1177/10775463251320258

[22] F Zouari, KB Saad, and M Benrejeb. Robust neural adaptive control for a class of uncertain nonlinear complex dynamical multivariable systems. International Review on Modelling and Simulations, 5(5):2075-2103, 2012.

[23] G. Rigatos, M. Abbaszadeh, B. Sari, P. Siano, G. Cuccurullo, and F. Zouari. Nonlinear optimal control for a gas compressor driven by an induction motor. Results in Control and Optimization, 11:100226, 2023. https://doi.org/10.1016/j.rico.2023.100226