Annals of Emerging Technologies in Computing (AETiC)
Vol. 10, No. 3, 2026

Research Article

A Residual Feature-Gated Variational Autoencoder with Transfer Learning for Economic Systemic Risk Exposure Prediction

Shiqiong Pan

College of Arts and Sciences, Boston University, Boston, USA
shiqiong@bu.edu

Received: 23 March 2026; Accepted: 27 June 2026; Published: 1 July 2026

Abstract: Accurate prediction of systemic economic risk is important for financial supervision and policy design, yet existing econometric and machine-learning models often suffer from limited data and unstable generalization. This paper proposes a two-step learning framework for predicting systemic risk exposure using a simulated economic resilience dataset. First, a Variational Autoencoder (VAE) is trained on real observations and used to generate 100,000 synthetic samples for data augmentation. Second, a residual neural network with a feature-gating mechanism is pretrained on the synthetic data and then fine-tuned on real samples through transfer learning. Experiments show that the proposed strategy leads to large performance gains. Compared with training from scratch, the two-stage model improves R^2 from 0.958 to 0.998, while reducing RMSE from 0.015 to 0.003 and MAE from 0.011 to 0.003. Classical models also benefit from augmentation: for SVR, R^2 increases from 0.018 to 0.090 after adding synthetic data. Feature screening based on mutual information highlights growth dynamics and risk-control indicators as the most influential drivers of systemic risk. Ablation studies further show that residual connections and feature gating are both essential, with the full model achieving the lowest error levels. These results demonstrate that generative data augmentation combined with transfer learning provides an effective and general solution for economic risk prediction under limited data conditions.

Keywords: ***Data Augmentation; Financial Stability; Systemic Risk Prediction; Transfer Learning; Variational Autoencoder***

1. Introduction

Understanding economic risk and resilience is an important task for governments, financial institutions, and policy makers [1, 2]. In complex market systems, capital allocation decisions, external shocks, and changing economic conditions can strongly affect long-term stability. Indicators that summarize overall risk exposure are therefore widely used to evaluate system vulnerability, stress resistance, and future performance. Reliable prediction of such indicators can help support planning, regulation, and early warning systems.

Traditionally, economic risk modeling has relied on statistical methods and econometric models, such as linear regression, factor analysis, and time-series forecasting [3, 4]. These approaches are often designed under strong assumptions about data distribution and variable relationships. While they are interpretable and efficient for small-scale problems, they may struggle when faced with large numbers of interacting factors, nonlinear effects, and complex feedback loops. In modern economic environments, risk is rarely driven by a single variable. Instead, it emerges from joint changes in market behavior, capital flows, volatility, and stress conditions. Capturing these patterns with fixed formulas can be difficult.

Shiqiong Pan, "A Residual Feature-Gated Variational Autoencoder with Transfer Learning for Economic Systemic Risk Exposure Prediction", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, 1 July 2026, Vol. 10, No. 3, pp. 32-49, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2026.03.003, Available: http://aetic.theiaer.org/archive/v10/v10n3/p3.html.

In recent years, machine learning methods have been increasingly applied to financial forecasting and risk assessment [5-7]. Neural networks, ensemble models, and representation learning techniques can automatically extract useful patterns from high-dimensional data and often achieve higher prediction accuracy than traditional approaches [8]. Transfer learning has also gained attention [9, 10], as it allows knowledge learned from one setting to be reused in another, reducing the need for large labeled datasets. At the same time, unsupervised or semi-supervised models such as autoencoders have been explored to learn compact representations of economic systems before performing downstream prediction tasks [11].

Despite recent progress, two major challenges remain in economic risk prediction. First, prediction accuracy is often limited when models are applied to complex systems with many interacting variables and nonlinear patterns. Traditional approaches may fail to capture these effects, while deep models can become unstable or overfit when data are scarce. Second, many economic risk datasets are relatively small or expensive to obtain, which restricts the training of data-hungry models and reduces their ability to generalize across different market conditions. These two issues motivate the need for modeling frameworks that can achieve reliable prediction performance while making efficient use of limited data.

This study focuses on the prediction of a composite economic risk indicator, referred to as systemic risk exposure, using a simulated economic resilience and risk modeling dataset. To address the challenges mentioned above, this paper proposes a two-step learning framework. First, a Variational Autoencoder (VAE) is trained on the available real data and then used to generate additional synthetic samples for data augmentation. Second, an artificial neural network with residual connections and an attention-inspired feature-gating mechanism is pretrained on the synthetic dataset and then fine-tuned on the real training set through transfer learning. The residual structure improves training stability, while the gating module highlights informative features and reduces the influence of less relevant inputs.

The main contributions of this study can be summarized as follows:

1) This paper proposes a two-stage learning framework for predicting systemic risk exposure, which integrates data augmentation based on Variational Autoencoder (VAE) and transfer learning to enhance the model's performance under conditions of limited data.

2) This paper develops a residual neural network with a feature gating mechanism to enhance training stability and selectively highlight the economically informative variables.

3) Compared with the baseline model, the proposed framework demonstrates superior predictive performance, indicating that it can effectively capture the complex nonlinear relationships within the economic system.

4) From an application perspective, this study provides a practical data-driven approach for improving risk monitoring, early warning systems, and macroprudential analysis in situations where data is scarce.

Furthermore, the concept of systemic risk exposure considered in this study is closely related to mature financial risk indicators such as CoVaR and Marginal Expected Loss (MES), which are widely used to quantify the contribution of individual entities to the overall systemic risk. From a theoretical perspective, the evolution of systemic risk can also be explained through mechanisms such as financial contagion and macroprudential dynamics, where shocks spread among interrelated entities and exacerbate overall instability. Therefore, the proposed data-driven framework can be regarded as a complementary approach, which can learn these complex interactions from data and conducting empirical approximations of potential economic risk transmission processes. The innovation of this study is integrating the synthesis data generation based on VAE, residual feature gated prediction, and transfer learning into a unified two-stage framework, which is used for systematic risk exposure prediction in the case of limited data. Unlike traditional models that directly train the predictor on the original dataset, the proposed framework first expands the data distribution through generative learning, and then transfers the learned information to the final prediction model.

2. Related Works

2.1. Modeling Approaches for Economic Risk Forecasting

Recent research on economic risk forecasting has increasingly applied machine learning and hybrid approaches to improve prediction performance for complex economic and financial systems. For example,

Wang *et al.* studied systemic risk prediction using several machine learning methods showed that incorporating network measures such as market connectedness can improve prediction accuracy of systemic risk indicators [12]. Tang *et al.* demonstrated that interpretable machine learning models can effectively forecast systemic financial stress and they identified key predictors such as stock–bond correlation and valuation risk that strongly influence systemic risk forecasts [13]. Other work has investigated hybrid econometric–machine learning frameworks for financial risk forecasting, finding that combinations like ARIMA–LSTM and GARCH–XGBoost can outperform traditional methods on macroeconomic and market volatility prediction tasks [14]. In addition, recent studies have proposed novel model architectures that integrate deep learning components to capture both spatial and temporal patterns in systemic risk data [15]. Finally, Srour explored forecasting systemic risk measures like ΔCoVaR and MES in the European banking industry using ANNs, support vector machines, and AR-GARCH models, showing the utility of machine learning tools in predicting risk contribution and exposure [16].

However, most existing studies rely on limited historical samples and assume that enough data are available to train complex models. In real economic systems, risk datasets are often small and unstable, which can reduce prediction accuracy and robustness. Current work mainly improves model structures, while less attention is paid to data augmentation and transfer learning strategies for economic risk forecasting. In particular, the use of generative models to expand training data for systemic risk exposure prediction remains limited.

2.2. Machine Learning Models-Based Augmentation Approaches in Financial Forecasting

Recent work has also explored the use of generative models to address data scarcity in financial forecasting. Some studies apply Generative Adversarial Networks (GANs) to augment financial time series data, showing that GAN-generated samples can help improve prediction models by providing additional training examples that mimic real market behavior and reduce noise effects. For example, researchers have used GAN variants such as TimeGAN and Wasserstein GAN to generate synthetic financial series that improve forecasting accuracy for market volatility and return trends over using only real data [17]. However, GAN-based approaches often face challenges such as training instability and difficulty capturing extreme events, which can limit their usefulness in risk estimation. To deal with these issues, Variational Autoencoders (VAEs) have been studied as an alternative generative model that is simpler to train and can produce smooth synthetic data. A unified evaluation of generative models in financial applications found that VAEs tend to provide more stable training outcomes and can still capture realistic temporal dynamics, making them useful for scenarios where stability is crucial [18].

Despite these advances, the application of generative augmentation specifically for economic systemic risk exposure prediction remains limited, and previous work has rarely integrated generative data enhancement with transfer learning pipelines to further boost model generalization under limited data conditions.

3. Optimized Variational Autoencoder with Transfer Learning for Economic Risk Exposure Prediction

3.1. Overview of the Developed Framework

A general illustration of the developed framework is provided in Figure 1. The first step focuses on data augmentation using an optimized VAE. The VAE is trained on the available real samples to learn the underlying distribution of economic states and then generates additional synthetic observations to expand the training set. Compared with a standard VAE, the proposed model incorporates residual connections and a lightweight feature-gating mechanism in both the encoder and decoder. The residual design helps stabilize training and allows deeper structures to be used, while the gating module adaptively reweights input features, emphasizing informative variables and reducing the impact of less relevant ones. In the second step, the synthetic data produced by the VAE are used to pretrain a prediction network based on residual multilayer perceptrons with feature gating. The pretrained model is then fine-tuned on the real training data through transfer learning, enabling it to adapt to the target distribution while retaining useful knowledge learned from the augmented dataset.

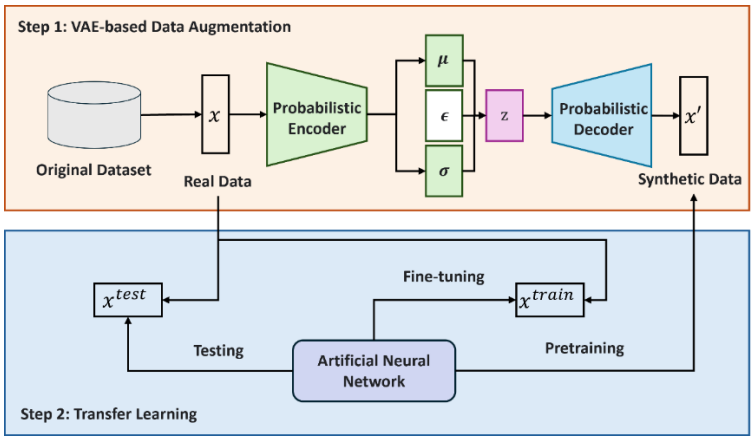

Figure 1. The proposed framework for economic risk exposure prediction

3.2. Detailed Architecture of the Two-Stage Learning Framework

Figure 2 presents the detailed architecture of the proposed two-stage learning framework, and the remainder of this section describes each module in turn.

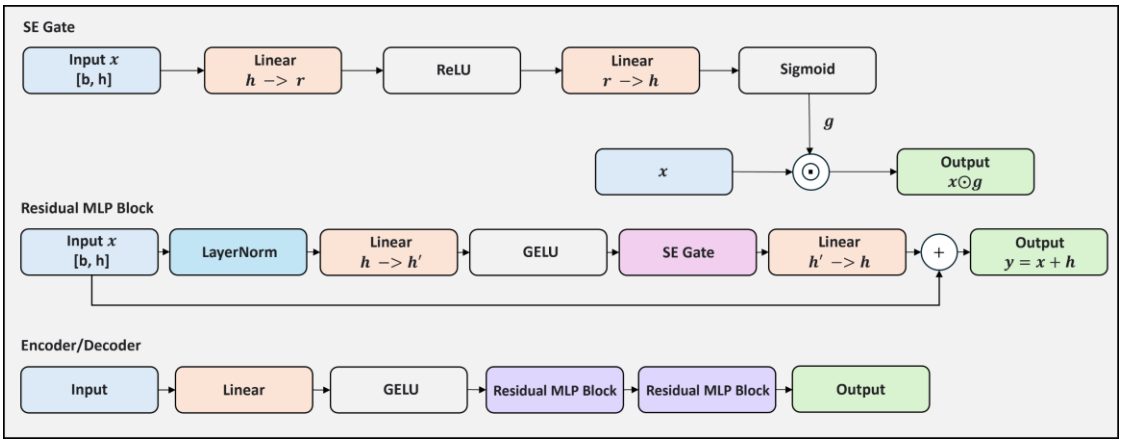

Figure 2. Detailed modules of the proposed framework

3.2.1. Feature-Gated Module (SE Gate)

To adaptively emphasize informative variables and suppress less relevant ones, a lightweight feature-gating module is introduced. Feature gating is a mechanism that assigns adaptive importance weights to each input variable. It is inspired by squeeze-and-excitation structures originally developed for deep neural networks [19]. Economic and financial indicators often show heterogeneous importance, and some variables may become noisy or redundant under different market regimes. This design keeps the model interpretable and stable for tabular economic data while allowing it to focus on the most relevant predictors of systemic risk exposure.

Given an input vector

$$\mathrm{x} \in \mathbb{R}^{H} \tag{1}$$

The gate first projects it into a low-dimensional space:

$$\mathrm{u} = \mathrm{ReLU}\,(\mathrm{W}_1\mathrm{x}+\mathrm{b}_1) \tag{2}$$

where $\mathrm{W}_1 \in \mathbb{R}^{r\times H}$ and

$$r = \max\,(1, \left\lfloor \frac{H}{\text{reduction}} \right\rfloor) \tag{3}$$

The reduced features are then mapped back to the original dimension:

$$\mathrm{g} = \sigma\,(\mathrm{W}_2\mathrm{u}+\mathrm{b}_2) \tag{4}$$

With $\mathrm{W}_2 \in \mathbb{R}^{H\times r}$ and sigmoid activation $\sigma(\cdot)$.

Finally, the gated output is obtained by element-wise multiplication:

$$\tilde{\mathrm{x}} = \mathrm{x} \odot \mathrm{g} \tag{5}$$

where $x \in \mathbb{R}^H$ is the input feature vector, and $H$is the number of input features. $u \in \mathbb{R}^r$represents the reduced feature representation after the first linear transformation and ReLU activation. r is the reduced hidden dimension, which is determined by the reduction ratio. $W_1 \in \mathbb{R}^{r \times H}$and $b_1 \in \mathbb{R}^r$are the weight matrix and bias vector of the first linear layer, while $W_2 \in \mathbb{R}^{H \times r}$and $b_2 \in \mathbb{R}^H$ are those of the second linear layer. $\sigma(\cdot)$ is the sigmoid activation function, and $g \in \mathbb{R}^H$ is the learned feature-gating vector. The symbol $\odot$ is element-wise multiplication, and $\tilde{x}$ is the reweighted output feature vector.

This mechanism performs feature-wise reweighting rather than cross-feature attention, making it simple and stable for tabular economic data.

3.2.2. Residual MLP Block

Residual learning is a widely used technique in deep neural networks that introduces shortcut connections between layers [20, 21]. Instead of forcing each block to learn a full transformation, the residual formulation allows the network to focus on modeling incremental refinements over the input. This greatly stabilizes optimization and alleviates degradation problems when networks become deep. For economic datasets with limited observations, deep multilayer perceptrons often overfit or suffer from unstable training. By embedding residual connections together with feature gating, the proposed block combines stable signal propagation with adaptive feature selection, forming a simple but effective nonlinear mapping for limited economic data.

Based on the gating module, a residual multilayer perceptron block is constructed. Given an input $x \in \mathbb{R}^H$, the block computes:

$$h_0 = \mathrm{LayerNorm}(x) \tag{6}$$

$$h_1 = \mathrm{Dropout}\left(\mathrm{GELU}(W_a h_0 + b_a)\right) \tag{7}$$

$$h_2 = \mathrm{SEGate}(h_1) \tag{8}$$

$$h_3 = \mathrm{Dropout}(W_b h_2 + b_b) \tag{9}$$

where $W_a, W_b \in \mathbb{R}^{H \times H}$.

A skip connection is finally applied:

$$y = x + h_3 \tag{10}$$

where $x \in \mathbb{R}^H$ is the input vector of the residual MLP block, and H is the feature dimension. h_0, h_1, h_2, and h_3 is the intermediate hidden representations generated by layer normalization, linear transformation, feature gating, and dropout operations, respectively. $W_a, W_b \in \mathbb{R}^{H \times H}$ are learnable weight matrices, and $b_a, b_b \in \mathbb{R}^H$ are learnable bias vectors. $\mathrm{LayerNorm}(\cdot)$, $\mathrm{GELU}(\cdot)$, $\mathrm{Dropout}(\cdot)$ and $\mathrm{SEGate}(\cdot)$ are layer normalization, GELU activation, dropout regularization and the feature-gating operation, respectively. y is the final output of the residual block after adding the skip connection.

This residual design improves gradient flow and stabilizes training when stacking multiple blocks.

3.2.3. Optimized Variational Autoencoder

A VAE is a probabilistic generative model that learns a low-dimensional latent representation of data and can generate new samples by sampling from this latent space [22-24]. VAEs are commonly used for data synthesis, representation vlearning, and uncertainty-aware modeling, which makes them well suited for economic systems characterized by noise and stochasticity. In this work, the VAE is optimized for tabular financial variables by replacing standard multilayer perceptrons with residual feature-gated blocks in both the encoder and decoder. This modification improves training stability and enables the latent variables to capture meaningful economic structures rather than noise.

The proposed VAE consists of an encoder and a decoder built from stacked residual blocks.

Encoder. For an input sample $x \in \mathbb{R}^D$, the encoder first maps it into the hidden space:

$$h = \mathrm{GELU}(W_e x + b_e) \tag{11}$$

followed by N_b residual blocks and a final layer normalization.

Two linear heads then produce the mean and log-variance:

$$\mu = W_\mu h, \log \sigma^2 = W_\sigma h \tag{12}$$

where $\mu, \sigma \in \mathbb{R}^Z$.

Reparameterization. A latent vector is sampled using:

$$z = \mu + \sigma \odot \epsilon, \epsilon \sim \mathcal{N}(0, I) \tag{13}$$

Decoder. The decoder maps z back to the data space:

$$\mathrm{h}' = \mathrm{GELU}(\mathrm{W}_d \mathrm{z} + \mathrm{b}_d) \quad (14)$$

followed by residual blocks and normalization, and outputs:

$$\hat{\mathrm{x}} = \mathrm{W}_o \mathrm{h}' + \mathrm{b}_o \quad (15)$$

In experiments, the hidden dimension is set to $H = 128$, latent dimension to $Z = 12$, and two residual blocks are used in both encoder and decoder.

where $\mathrm{x} \in \mathbb{R}^D$ is the input sample and D is the input feature dimension. h and h′ are the hidden representations in the encoder and decoder, respectively. N_b is the number of residual blocks. μ and σ are the mean and standard deviation vectors of the latent distribution, and Z is the latent dimension. $\mathrm{z} \in \mathbb{R}^Z$ is the sampled latent vector, while $\epsilon \sim \mathcal{N}(0, \mathrm{I})$ is a random noise vector sampled from a standard normal distribution. $\hat{\mathrm{x}}$ is the reconstructed output sample. The matrices W_e, W_μ, W_σ, W_d, and W_o, together with the bias terms b_e, b_d, and b_o, are learnable parameters of the encoder and decoder.

3.2.4. Prediction Network and Transfer Learning

Transfer learning refers to the practice of pretraining a model on one dataset or task and then adapting it to a related target problem [25, 26]. It is particularly useful when labeled data in the target domain are scarce or expensive to obtain, which is often the case for systemic risk studies relying on simulated or historical crisis data. In the proposed framework, the prediction network is first trained on VAE-generated synthetic samples to learn general economic patterns. It is then fine-tuned using real observations so that the model adapts to the true data distribution. This two-stage procedure improves generalization while keeping the final predictor simple and resistant to overfitting.

For systemic risk exposure prediction, a shallow multilayer perceptron is adopted:

$$\mathrm{h} = \mathrm{ReLU}(\mathrm{W}_1 \mathrm{x} + \mathrm{b}_1), \hat{y} = \mathrm{W}_2 \mathrm{h} + \mathrm{b}_2 \quad (16)$$

where the hidden layer contains 4 neurons.

The prediction model is first pretrained on VAE-generated synthetic samples and then fine-tuned on the real training data. This transfer learning strategy allows the model to exploit the enlarged dataset while adapting to the true distribution.

4. Experimental Setup

4.1. Dataset Description

The experiments are conducted on the Economic Resilience & Risk Modeling Dataset [27], a public benchmark designed for AI-based financial forecasting and risk analysis. The dataset simulates the behavior of 10,000 economic agents operating over multiple economic cycles under changing market conditions and regulatory environments. Therefore, the original dataset used for model development has 1,000 records. Although this dataset is simulated, it is constructed based on established economic principles and incorporates real market dynamics such as capital allocation, volatility, and stress responses to improve its practical relevance. Therefore, the patterns learned can provide useful approximations for the real-world financial system and offer generalizable insights for systemic risk modeling.

Two representative capital allocation strategies are included in the dataset. The first group focuses on aggressive expansion with higher risk exposure, while the second emphasizes sustainable efficiency and resilience to adverse shocks. For each agent, the dataset records a collection of financial and macroeconomic indicators, including capital dynamics, growth and efficiency measures, liquidity conditions, systemic stress responses, and survival-related metrics. All variables are stored in tabular form, where each row corresponds to the aggregated performance profile of one agent. This structure makes the dataset suitable for studying systemic risk forecasting, capital allocation analysis, and stress-testing scenarios in a controlled simulation environment.

4.2. Prediction Target and Feature Selection

The goal of this study is to predict WeightedRiskExposure, which is a composite indicator reflecting the overall level of systemic financial risk faced by an economic agent. This variable summarizes the joint effects of volatility, macroeconomic stress sensitivity, and exposure to adverse market dynamics, and therefore serves as the main dependent variable throughout the experiments.

An entity identifier (EntityID) is excluded from modeling because it does not contain predictive economic information. In addition, several variables, namely CompetitivePressure, AdaptationIndex, RelativeAdaptationIndex, MarketAdaptability, and InnovationFactor, are removed from the input feature set based on a preliminary feature importance analysis conducted on the training data. This analysis indicates that these variables contribute marginally to the prediction of WeightedRiskExposure.

To avoid unnecessary model complexity and reduce the influence of weakly informative inputs, these features are therefore excluded from subsequent modeling. A detailed discussion of the feature importance results is provided in the experimental results section. The categorical variable ModelType, which represents the strategic group of each agent, is retained as an explanatory factor and encoded during preprocessing.

4.3. Data Preprocessing

Before model training, several preprocessing steps are applied to ensure numerical stability and consistency across real and synthetic samples. All preprocessing operations are strictly carried out using only the training data, and then the same transformation parameters are applied to the validation set and the test set to prevent data leakage.

All continuous variables are converted to numeric form, and infinite values are replaced using the median computed from the training data. Each numerical feature is then standardized using z-score normalization:

$$x' = \frac{x-\mu}{\sigma} \tag{17}$$

where x is the original numerical feature value, x' is the standardized feature value, and μ and σ denote the mean and standard deviation estimated from the training subset.

The categorical variable ModelType is transformed into integer labels based on the mapping learned from the training data only, preventing information leakage into validation or test sets.

For the VAE-based data augmentation stage, standardized numerical features and encoded categorical features are concatenated to form the final input matrix. The resulting feature dimension defines the input size of the encoder and decoder networks.

4.4. Dataset Splits

The dataset is first divided into training and testing subsets using an 80/20 split at the agent level. The training portion is further divided into training and validation sets, with 20% reserved for hyperparameter tuning and early stopping.

For the VAE training stage, only real training samples are used to learn the underlying data distribution. These samples are provided to the model through mini-batch loaders with a batch size of 256. After training, the VAE is used to generate a large synthetic dataset consisting of 100,000 samples, which is employed for data augmentation.

The prediction neural network is then trained under two settings. First, it is pretrained using the VAE-generated synthetic data. Second, it is fine-tuned on the real training subset and evaluated on the held-out test set.

4.5. Training Settings and Hyperparameters

All models are implemented in PyTorch and trained using the Adam optimizer. Mini-batch training is adopted for both the generative and prediction stages, with a batch size of 256. Mean squared error is used as the reconstruction loss for the VAE and as the regression loss for the prediction network. When available, GPU acceleration is used for all experiments. To ensure reproducibility, a fixed random seed of 42 was used for all experiments across all libraries.

4.5.1. VAE Training Configuration

The variational autoencoder is trained only on real training samples in order to learn the underlying data distribution before generating synthetic observations. The dataset is provided through shuffled mini-batches, while a separate validation set is used for monitoring convergence.

The VAE is optimized using Adam with a learning rate of 10^{-3}. Training is performed for 300 epochs. The objective function consists of a reconstruction term and a Kullback–Leibler divergence regularization term:

$$\mathcal{L}_{\text{VAE}} = \underbrace{\| \hat{x} - x \|_2^2}_{\text{reconstruction}} + \beta \underbrace{(-\frac{1}{2}\, \mathbb{E}[1 + \log \sigma^2 - \mu^2 - \sigma^2])}_{\text{KL divergence}} \quad (18)$$

where β is set to 1.0 in all experiments. The reconstruction term is computed using mean squared error, while the KL term enforces a standard normal prior on the latent variables.

The complete set of training hyperparameters is summarized in Table 1. Due to the relatively small size of the actual training dataset and the limited dimension of the input features, the hidden layer is set to have 4 neurons.

Table 1. Training Hyperparameters for the VAE

Hyperparameter	Optimal Value
Optimizer	Adam
Learning rate	1×10^{-3}
Batch size	256
Training epochs	300
Reconstruction loss	MSE
KL weight $\boldsymbol{\beta}$	1.0
Hidden dimension	128
Latent dimension	12
Residual blocks	2

4.5.2. ANN Transfer Learning Configuration

The prediction model is trained in two stages following the proposed transfer-learning strategy.

In the first stage, the residual MLP regressor is pretrained on the VAE-generated synthetic dataset using a learning rate of 10^{-3}. No validation set is used during this phase, as the goal is to initialize the network with broad structural patterns learned from the augmented data.

In the second stage, the pretrained model is fine-tuned on the real training data for up to 10 epochs, with a reduced learning rate of 5×10^{-4}. A validation subset is employed for early stopping. Weight decay of 1×10^{-4} is applied in both stages to reduce overfitting.

The regression loss is defined as

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \quad (19)$$

where y_i and $\hat{y}_i$ denote the true and predicted systemic risk exposure values.

The complete set of training hyperparameters is summarized in Table 2.

Table 2. Training Hyperparameters for the ANN

Hyperparameter	Optimal Value
Optimizer	Adam
Fine-tuning epochs (real)	10
Learning rate (pretrain)	1×10^{-3}
Learning rate (fine-tune)	5×10^{-4}
Batch size	256
Weight decay	1×10^{-4}
Loss function	MSE
Hidden neurons	4

5. Empirical Results on Systemic Risk Exposure Prediction

5.1. Feature Screening Based on Mutual Information

To better understand which variables are most relevant for predicting WeightedRiskExposure, a feature importance analysis based on mutual information is conducted. Mutual information measures the statistical dependency between each input variable and the target, allowing both linear and nonlinear relationships to be captured. A higher mutual information value indicates that a feature provides more useful information for explaining variations in the risk exposure indicator.

Figure 3 reports the top-ranked variables according to their mutual information scores. The results show that factors related to market growth and risk control play a dominant role. In particular, ReinforcementSignal and GrowthRate exhibit the highest scores, suggesting that adaptive investment feedback and expansion dynamics strongly influence systemic risk exposure. Variables such as RiskManagementScore, Efficiency, and SustainabilityScore also appear among the most informative inputs, highlighting the importance of operational discipline and long-term stability in shaping aggregate risk outcomes.

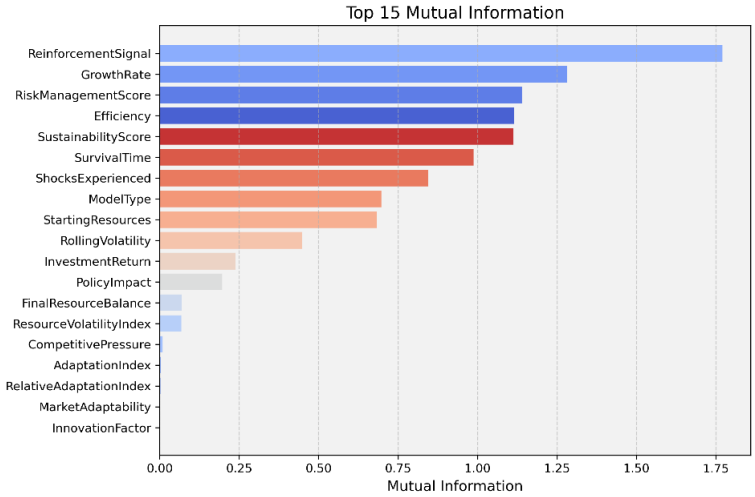

Figure 3. The importance of each variable based on mutual information

In addition, indicators related to survival and market stress, including SurvivalTime and ShocksExperienced, provide meaningful explanatory power. The categorical variable ModelType, which represents the strategic class of each agent, is also moderately informative, implying that different capital allocation styles are associated with distinct risk profiles.

Several variables, namely CompetitivePressure, AdaptationIndex, RelativeAdaptationIndex, MarketAdaptability, and InnovationFactor, receive near-zero mutual information scores and therefore contribute little to predicting WeightedRiskExposure in this dataset. Based on this analysis, these variables are excluded from the input feature set in subsequent experiments. As discussed later in the results section, this feature screening step helps the models focus on financially observable signals while avoiding variables that provide limited predictive value in practice.

5.2. Distribution Comparison Between Real and Synthetic Samples

Before analyzing the distribution of the generated samples, this paper first examines the training behavior of the proposed VAE. Figure 4 reports the evolution of the training and validation losses over epochs. Both curves decrease rapidly in the early stage and gradually converge, indicating that the model successfully learns the underlying data distribution without severe overfitting. The close gap between the two curves suggests stable optimization and good generalization, which provides a reliable basis for subsequent data generation and transfer learning experiments.

To evaluate whether the VAE-generated samples realistically reflect the structure of the original dataset, this study visualizes the distributions of real and synthetic data using Principal Component Analysis (PCA) and T-distributed Stochastic Neighbor Embedding (t-SNE) [28, 29]. Both techniques project high-dimensional feature vectors into two-dimensional spaces for intuitive inspection. PCA provides a linear projection that preserves major variance directions, while t-SNE focuses on preserving local neighborhood relationships and is widely used to assess distributional similarity in generative modeling.

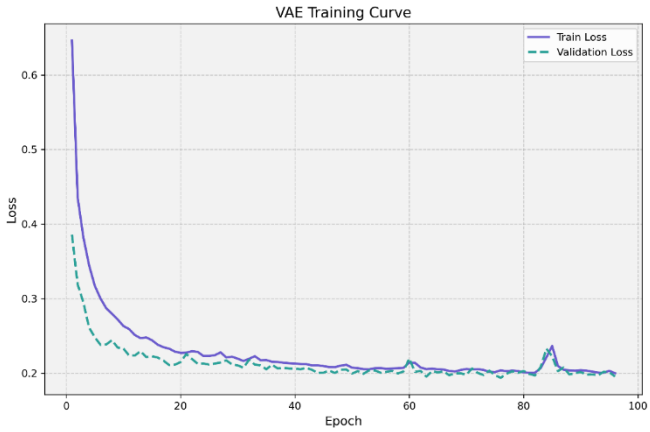

Figure 4. Comparison of real and synthetic data distributions using PCA and t-SNE projections

Figure 5 presents the comparison between real and synthetic samples under these two projections. In the PCA visualization, the synthetic data largely follows the same curved manifold as the real observations, indicating that the dominant variance patterns of the economic variables are well captured by the proposed VAE. Although some dispersion is visible at the tails of the distribution, the overall alignment suggests that the generated samples reproduce the main global structure of the original feature space. This observation is further supported by the calculated Wasserstein distance in the projection space, which has an average value of 16.91 across the two projected dimensions. Although there are still negligible differences, the results indicate that the generated samples retain the overall distribution characteristics of the real data, while allowing for moderate variability, thereby enhancing the generalization ability of the model.

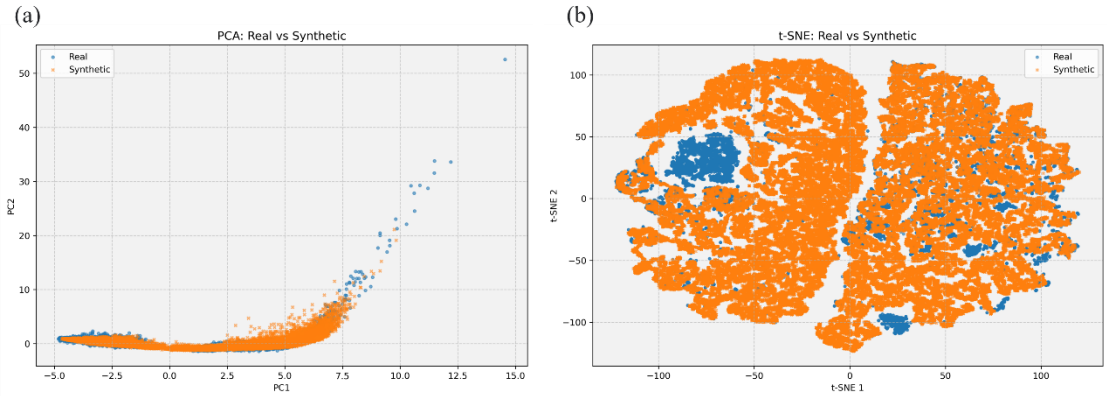

Figure 5. Comparison of real and synthetic data distributions using PCA and t-SNE projections

The t-SNE plot further confirms this observation. Real and synthetic samples exhibit strong overlap across most regions, and the generated points cover nearly all dense clusters formed by the real data. This indicates that the VAE is able to reproduce fine-grained local structures and heterogeneous agent behaviors present in the simulation. Importantly, the synthetic data do not collapse into a small number of modes, suggesting that the model avoids severe mode collapse and maintains diversity in the generated economic states.

These visual results demonstrate that the optimized VAE, enhanced with residual blocks and feature-gating modules, is capable of learning a realistic approximation of the underlying data distribution. The close match between real and synthetic samples provides empirical support for using the generated dataset to pretrain the prediction network in the subsequent transfer learning stage.

5.3. Prediction Performance Comparison with and without VAE-Based Augmentation

Figure 6 compares prediction results obtained when training the neural network from scratch and when applying the proposed VAE-based data augmentation and transfer learning strategy. Scatter plots of predicted versus true values show that the model trained from scratch produces a wider spread around the diagonal line, indicating noticeable prediction errors, especially for low-risk and mid-risk regions. In

contrast, the two-stage model produces predictions that align much more closely with the ideal diagonal line, suggesting substantially improved accuracy and consistency.

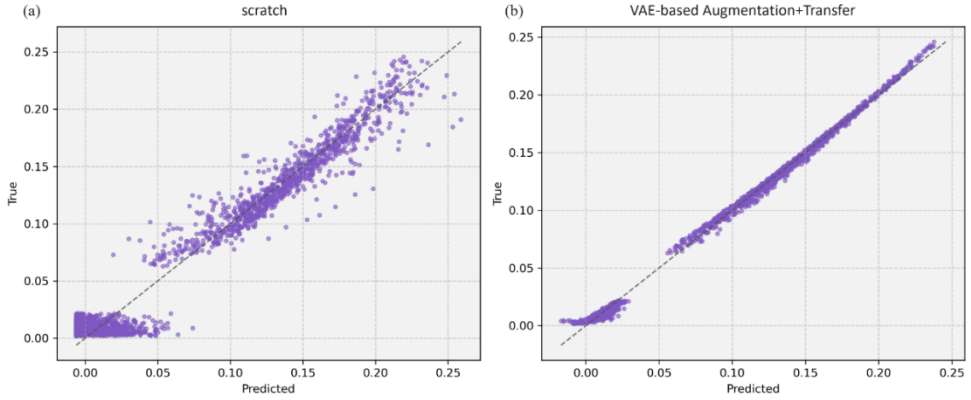

Figure 6. Predicted versus true values of WeightedRiskExposure for models trained from scratch and with the proposed VAE-based augmentation and transfer learning strategy

This improvement can be attributed to the transfer learning strategy. This strategy utilizes the data generated by the VAE for pre-training, enabling the model to learn a more comprehensive representation of the underlying data distribution to provide a better initialization for subsequent fine-tuning on real data. As a result, the model can converge to a more stable solution and achieve better generalization performance.

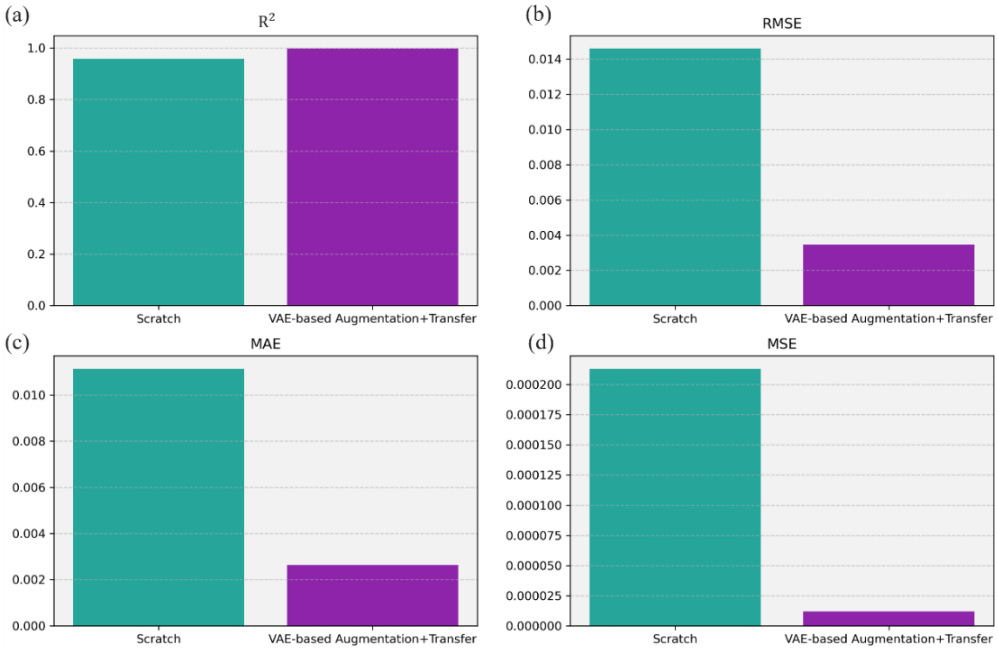

Figure 7. Comparison of prediction performance between baseline training from scratch and the proposed two-stage framework in terms of R^2, RMSE, MAE, and MSE

Quantitative results further confirm this observation. As illustrated in Figure 7, the proposed approach achieves a higher coefficient of determination (R^2) and substantially lower RMSE, MAE, and MSE compared with the baseline model trained only on real data. Specifically, the R^2 value obtained from training from scratch was 0.958, while the value was increased to 0.998 by using data augmentation and transfer strategies based on VAE. At the same time, the RMSE decreased from 0.015 to 0.003, the MAE from 0.011 to 0.003, and the MSE from 2.13×10^{-4} to 1.20×10^{-5}. All these indicators are calculated on the original scale of the target variable after inverse transformation, rather than on the normalized target values. Therefore, the reported RMSE, MAE and MSE directly reflect the prediction errors of the weighted risk exposure model on its original data scale. These quantitative improvements also indirectly prove the quality of the generated synthetic data, as the enhanced samples help the model learn useful patterns and improve its prediction accuracy on the independent test set. The improvement in R^2 indicates that the two-stage framework explains a much larger portion of the variance in the target variable

WeightedRiskExposure, while the reductions in error metrics demonstrate more accurate point-wise predictions. In addition, a paired t-test was conducted on the prediction errors of the two models. The results showed that the improvement was statistically significant ($p < 0.05$), confirming that the performance enhancement was not due to random fluctuations.

These gains can be attributed to two complementary factors. First, the VAE-generated synthetic samples expand the effective training set and expose the prediction network to a broader range of economic states, which is especially valuable when the number of real observations is limited. Second, pretraining on the synthetic data allows the model to learn general patterns before being fine-tuned on the real dataset, leading to faster convergence and better generalization.

To further examine whether the proposed VAE-based data augmentation strategy is beneficial beyond neural networks, two widely used classical regression models, namely K-nearest Neighbors (KNN), Support Vector Regression (SVR) and XGBoost [30, 31], are considered as additional baselines. These models are trained under two settings: using only the original dataset, and using the combination of original data and VAE-generated synthetic samples.

Figure 8 presents the predicted-versus-true scatter plots for KNN, SVR and XGboost under both settings. When trained on the original data alone, KNN already shows strong performance, while SVR exhibits noticeable bias and dispersion, especially for medium and high values of WeightedRiskExposure. For XGBoost, the predicted values are also closely distributed along the diagonal, indicating its stable performance. Compared with KNN and SVR, XGBoost shows no significant change after data augmentation because its predicted values are already in good agreement with the true values under the original settings. After incorporating the synthetic data generated by the VAE, all models display tighter alignment with the diagonal reference line, indicating improved predictive accuracy and reduced variance. The improvement is particularly clear for SVR, which benefits from the enlarged training set and achieves more stable predictions across the full range of target value.

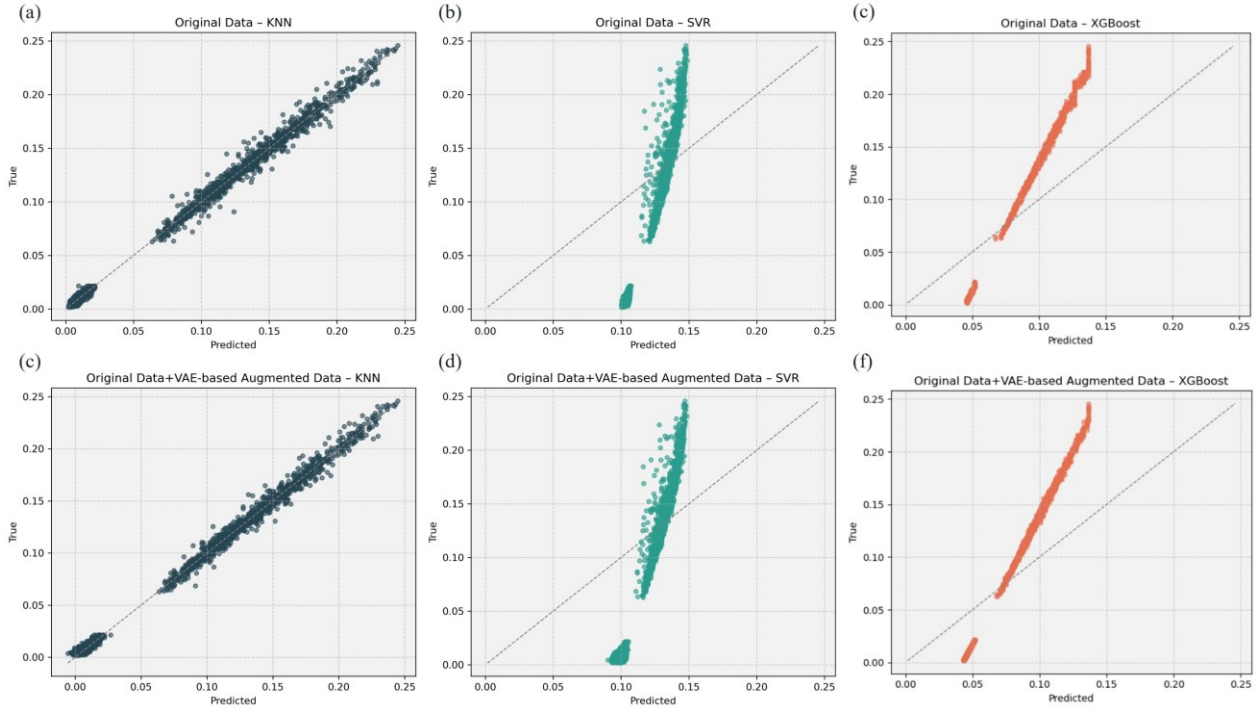

Figure 8. Predicted versus true WeightedRiskExposure for KNN, SVR, XGBoost trained on the original dataset and on the combination of original and VAE-augmented data

After incorporating the synthetic data generated by the VAE, all models show consistent improvement. For KNN, R^2 slightly increases to 0.996, while RMSE and MAE decrease to 0.004 and 0.003, respectively. Although the gain is moderate due to its already strong baseline performance, the augmented data still provide a measurable benefit. For SVR, the improvement is more pronounced. The R^2 increases from 0.018 to 0.090, while RMSE drops from 0.071 to 0.068 and MAE decreases from 0.061 to 0.058. For XGBoost, the performance remains relatively stable after incorporating the generated data. The R^2 value shows a slight increase from 0.648 to 0.6483. These results suggest that the enlarged training set

helps SVR better capture the nonlinear structure of the economic system and improves its stability across the full range of target values.

The quantitative comparison in Figure 9 further confirms these observations. For both KNN, SVR and XGBoost, VAE-based augmentation leads to higher R^2 values and lower RMSE, MAE, and MSE compared with training on the original data alone. Although KNN performs well in both cases, the augmented data still provides a small but consistent gain. In contrast, SVR shows a more noticeable improvement, suggesting that data augmentation is especially useful for models that are sensitive to training sample size and data coverage.

These results indicate that the proposed augmentation strategy is not limited to deep learning models, but can also enhance the performance of classical machine learning approaches. This strengthens the general applicability of the two-stage framework and highlights the usefulness of VAE-generated samples for economic risk prediction tasks under limited data conditions.

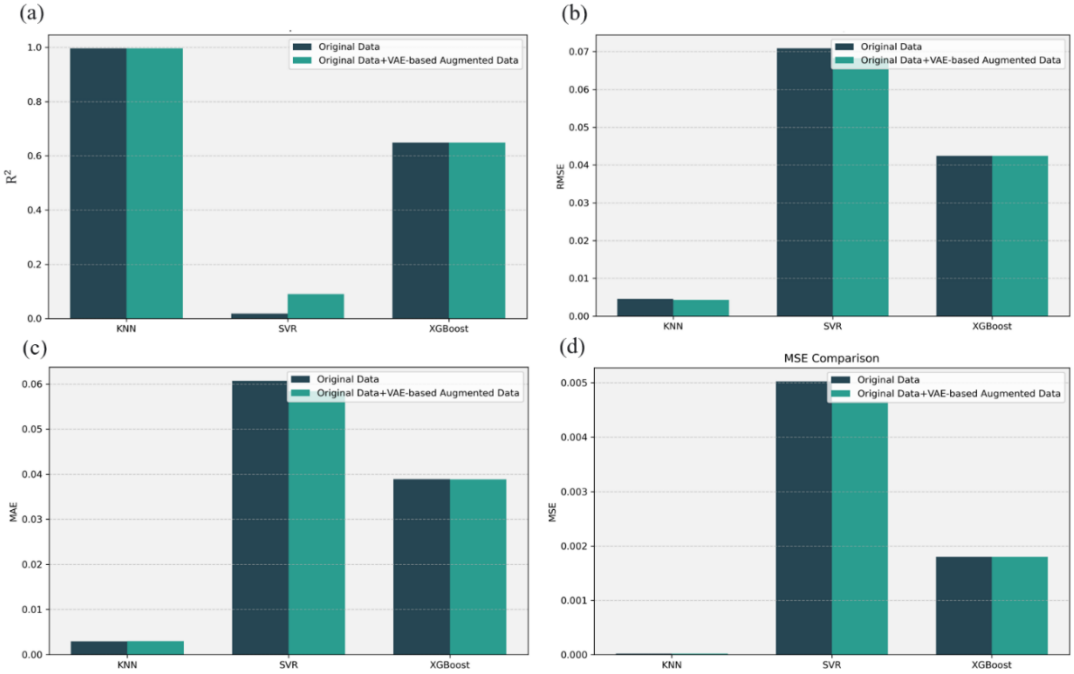

Figure 9. Performance comparison of KNN, SVR, XGBoost with and without VAE-based data augmentation in terms of R^2, RMSE, MAE, and MSE

5.4. Hyperparameter Sensitivity Analysis of the Transfer Learning Model

To examine the influence of key training parameters on model performance, this study conducts a sensitivity analysis on the learning rate and weight decay during the fine-tuning stage of the transfer learning model. The prediction network is first pretrained on VAE-generated samples and then fine-tuned on real training data. Performance is evaluated on the test set using R^2, RMSE, MAE, and MSE.

Figure 10 shows the effect of the learning rate on prediction accuracy. As the learning rate increases from 1×10^{-4} to 5×10^{-4}, R^2 rises steadily, while RMSE, MAE, and MSE decrease. This trend indicates that larger learning rates within this range help the model adapt more effectively to the real-data distribution during fine-tuning. The best overall performance is achieved around 4×10^{-4} to 5×10^{-4}.

Figure 11 reports the results for different weight decay values. Moderate regularization leads to better performance, whereas overly small or large values reduce accuracy. In particular, weight decay around 1×10^{-4} provides the highest R^2 and relatively low error levels, suggesting a good trade-off between model flexibility and overfitting control.

Based on these observations from Figures 10 and Figure 11, the final model adopts a learning rate of 5×10^{-4} and a weight decay of 1×10^{-4} for all subsequent experiments.

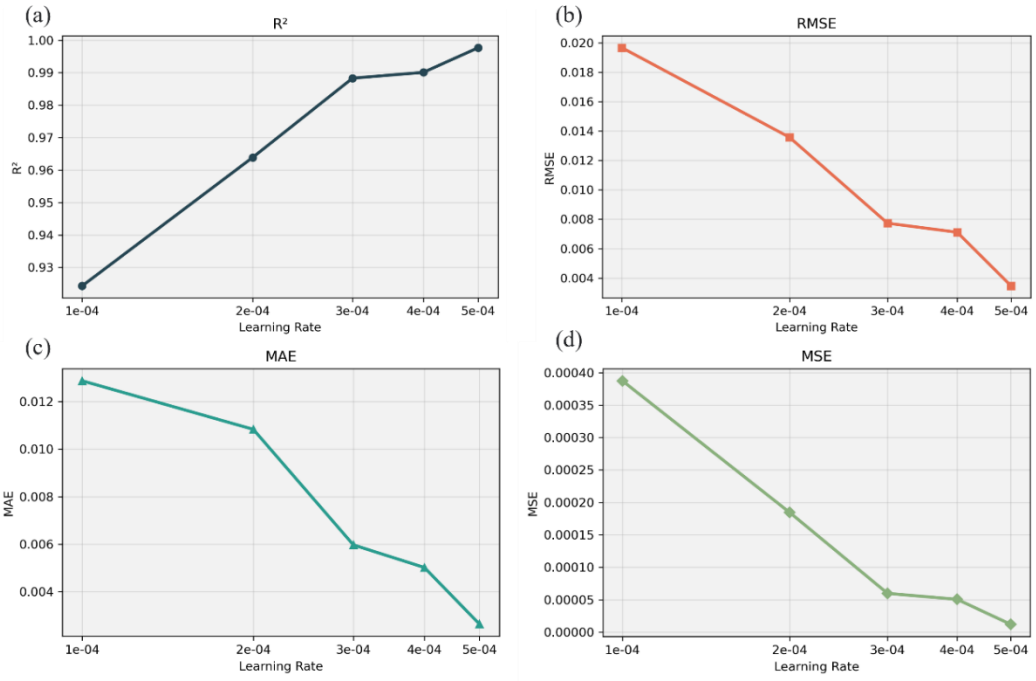

Figure 10. Sensitivity of prediction performance to the learning rate during transfer learning, evaluated using R^2, RMSE, MAE, and MSE

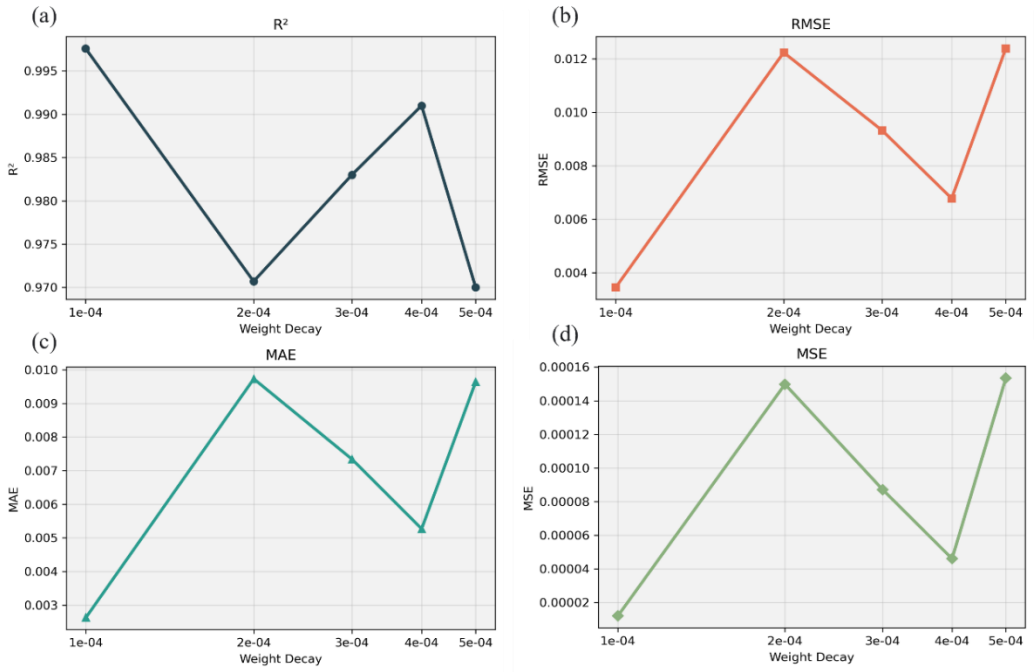

Figure 11. Effects of weight decay on transfer learning performance in terms of R^2, RMSE, MAE, and MSE

5.5. Ablation Study on Residual Connections and Feature Gating

To examine the role of key architectural components in the proposed prediction network, an ablation study is conducted focusing on residual connections and the feature gating module. Three model variants are evaluated: (1) a simplified version without both residual connections and feature gating, (2) a residual-only version where the gating mechanism is removed, and (3) the full model that integrates both components.

As reported in Table 3, removing both modules leads to the weakest performance, with R^2 dropping to 0.989 and RMSE increasing to 0.008. This indicates that these structures are crucial for effective transfer learning and stable optimization. When feature gating is excluded but residual connections are retained, performance improves notably (R^2 = 0.995), yet all error metrics remain higher than those of the full

model. This suggests that feature gating plays an important role in guiding the network to focus on informative economic variables during fine-tuning.

The complete architecture achieves the best results, reaching an R^2 of 0.998 and the lowest RMSE, MAE, and MSE. These findings demonstrate that residual connections and feature gating act in a complementary manner, jointly enhancing predictive accuracy and robustness for systemic risk exposure forecasting. To further evaluate the stability and generalization ability of the proposed model, we adopted a three-fold cross-validation strategy. The average performance was R^2 = 0.9977, RMSE = 0.0029, MAE = 0.0028, and MSE = 1.18×10^{-5}, which is chose to the results reported in Table 3. These results indicate that the proposed model achieves a high accuracy rate across different data set divisions, which verifies its robustness and reliability. Since the model has been evaluated on an independent test set and demonstrated consistent results in threefold cross-validation, the likelihood of severe overfitting is limited. Therefore, the high R^2 value primarily reflects the stable generalization capability of the proposed framework.

Table 3. Ablation study on residual connections and feature gating modules in the transfer learning model

Setting	R^2	RMSE	MAE	MSE
Without Residual & Feature Gating	0.989	0.008	0.005	5.82e−05
Without Feature Gating (Residual Only)	0.995	0.005	0.004	2.78e−05
Full Model (Residual + Feature Gating)	0.998	0.003	0.003	1.19e−05

5.6. Discussion

The empirical results reported in Sections 5.1–5.5 provide consistent evidence that the proposed two-stage learning framework, which combines VAE-based data augmentation with transfer learning, substantially improves systemic risk exposure prediction.

From a methodological perspective, the performance improvement can be attributed to the complementary effects of three components. The data augmentation based on Variational Autoencoder (VAE) expands the training distribution and reduces overfitting in the case of limited data. The residual connections stabilize the optimization process for learning complex nonlinear relationships, while the feature gating mechanism adaptively enhances the informative variables and suppresses noise. These components jointly enhance the representation learning and generalization performance.

From an economic perspective, the improvement in prediction accuracy indicates that this model is better able to capture the hidden interdependencies among financial entities and the nonlinear risk transmission mechanisms, which are the key characteristics of systemic risks. In actual financial risk management, more accurate prediction of systemic risk exposure helps regulatory authorities identify vulnerable institutions or market segments before the risks fully manifest. It can also support the early warning system by providing quantitative signals, thereby determining whether an entity is moving towards a higher-risk state. For financial institutions, such predictions can serve as a reference for capital allocation, portfolio adjustment, and stress testing design. Taken together, these results suggest several promising directions for applying the proposed framework in financial risk monitoring systems. In real-world settings, supervisory agencies and financial institutions often face sparse observations of extreme events and rapidly evolving market structures. VAE-based scenario generation combined with transfer learning offers a principled way to enrich training data, stress-test predictive models under rare but plausible conditions, and improve early-warning indicators of systemic instability. The framework could be integrated into macroprudential surveillance platforms or agent-based market simulators to support regulatory decision-making and portfolio risk management. For example, regulatory authorities can utilize the model outputs to identify institutions with higher marginal risk contributions, while financial institutions can use the prediction results to adjust their asset allocation and hedge against potential systemic shocks.

However, several limitations and challenges remain. First, although the VAE captures the empirical distribution of the available data, generative models may still underrepresent truly extreme tail events unless explicitly constrained or augmented with stress scenarios. Second, the quality of synthetic samples depends on the representativeness of the original dataset; structural breaks or regime shifts in real markets may reduce the usefulness of historical patterns learned by the VAE. Third, computational costs

associated with training generative models and performing two-stage optimization may limit real-time applications unless efficient updating strategies are developed. Finally, interpretability remains an important issue: while the framework improves predictive accuracy, further work is needed to link latent representations and synthetic scenarios to economically meaningful mechanisms that can be audited by regulators.

6. Conclusion

This paper presents a VAE-based data augmentation and transfer-learning framework for systemic risk exposure prediction. The proposed approach shows better performance than models trained only on real data under the current experimental setting. The neural predictor achieves an R^2 of 0.998 with RMSE of 0.003, compared with 0.958 and 0.015 when trained from scratch. Classical models also improve after augmentation, especially SVR, whose R^2 rises from 0.018 to 0.090. Ablation experiments confirm that both residual connections and feature-gating modules play key roles in achieving stable optimization and high accuracy.

These findings suggest that generative models can be used to expand limited economic datasets and strengthen downstream forecasting models in stress-testing and policy-analysis settings. It is important to note that this study is based on simulated data sets, so the results may not fully reflect the actual financial situation. Although the simulation framework allows for the controlled assessment of model behavior, it may also limit the external validity of the research results. Therefore, the general applicability of the proposed framework should be interpreted with caution before it is validated on real-world financial datasets. Several challenges remain, including modeling extreme tail events, handling regime shifts, and improving interpretability of synthetic scenarios, as well as incorporating evaluation metrics that better reflect financial relevance, such as risk ranking accuracy and tail-risk performance. In future work, we will use real-world financial datasets to validate the proposed framework, in order to further assess its practical applicability and robustness. Future research will also focus on stress-aware data generation, adaptive model updating under changing market conditions, and the connection between latent representations and economically meaningful mechanisms for regulatory use. In addition, more robust feature importance analysis methods, such as SHAP and permutation-based techniques will be used to further validate the relevance of selected variables.

CRediT Author Contribution Statement

Shiqiong Pan: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft, Writing – review & editing.

References

[1] Stéphane Hallegatte, "Economic resilience: definition and measurement", *World Bank Policy Research Working Paper 6852*, Print ISSN: 0198-5177, Online ISSN: 1640-0055, 1 May 2014, Vol. 1, No. 1, pp. 1-39, Published by World Bank, DOI: 10.1596/1813-9450-6852, Available: http://documents.worldbank.org/curated/en/350411468149663792.

[2] Bram Tucker and Donald R. Nelson, "What does economic anthropology have to contribute to studies of risk and resilience?", *Economic Anthropology*, Print ISSN: 2330-3635, Online ISSN: 2330-3643, 2 June 2017, Vol. 4, No. 2, pp. 161-172, Published by Wiley, DOI: 10.1002/sea2.12086, Available: https://anthrosource.onlinelibrary.wiley.com/doi/abs/10.1002/sea2.12085.

[3] Raisa Viktorovna Levkina, Ivan Ivanovych Kravchuk, Iryna Volodymyrivna Sakhno, Kateryna Mykhailivna Kramarenko and Anatolii Aanatoliiovych Shevchenko, "The economic-mathematical model of risk analysis in agriculture in conditions of uncertainty", *Financial and Credit Activity Problems of Theory and Practice*, Print ISSN: 2311-9385, Online ISSN: 2311-9407, 1 September 2019, Vol. 3, No. 30, pp. 248-255, Published by KNUTE, DOI: 10.18371/fcaptp.v3i30.179560, Available: https://doaj.org/article/e2c757117a644a0b88019001f2f58e5a.

[4] Marc Nekhom and Arthur Wayne Barker, "Economic Risk Analysis Model", in *SPE Hydrocarbon Economics and Evaluation Symposium*, 11–13 February 1979, Dallas, Texas, SPE-7731-MS, ISBN: 978-1-55563-718-7, Vol. 1979, No. 1, pp. 1-10, Published by Society of Petroleum Engineers, DOI: 10.2118/7731-MS, Available: https://onepetro.org/SPEHEES/proceedings-abstract/79HEE/79HEE/SPE-7731-MS/134565.

[5] Akib Mashrur, Wei Luo, Nayyar Zaidi and Antonio Robles-Kelly, "Machine learning for financial risk management: a survey", *IEEE Access*, Print ISSN: 2169-3536, Online ISSN: 2169-3536, 1 January 2020, Vol. 8, No. 1,

pp. 203203-203223, Published by IEEE, DOI: 10.1109/ACCESS.2020.3036322, Available: https://ieeexplore.ieee.org/document/9249416.

[6] Shiyang Chen and Shaochen Ren, "AI-enabled Forecasting, Risk Assessment and Strategic Decision Making in Finance", *Frontiers in Business and Finance*, Print ISSN: 2813-0422, Online ISSN: 2813-0430, 1 April 2025, Vol. 2, No. 02, pp. 274-295, Published by Frontiers Media S.A., DOI: 10.71465/fbf397, Available: https://sprcopen.org/index.php/FBF/article/view/397.

[7] Oluwabusayo Adijat Bello, "Machine learning algorithms for credit risk assessment: an economic and financial analysis", *International Journal of Management*, Print ISSN: 0972-7668, Online ISSN: 0972-7676, 17 June 2024, Vol. 10, No. 1, pp. 109-133, Published by European – American Journals, DOI: 10.37745/ijmt.2013/vol10n1109133, Available: https://eajournals.org/ijmt/vol10-issue-1-2023/machine-learning-algorithms-for-credit-risk-assessment-an-economic-and-financial-analysis/.

[8] Boyu Zhang, Ji Xiang and Xin Wang, "Network representation learning with ensemble methods", *Neurocomputing*, Print ISSN: 0925-2312, Online ISSN: 1872-8286, 7 March 2020, Vol. 380, No. 1, pp. 141-149, Published by Elsevier B.V., DOI: 10.1016/j.neucom.2019.10.098, Available: https://www.sciencedirect.com/science/article/abs/pii/S0925231219315413.

[9] Yuhang Qiu, Yunze Hui, Pengxiang Zhao, Mengting Wang, Shirong Guo *et al.*, "The employment of domain adaptation strategy for improving the applicability of neural network-based coke quality prediction for smart cokemaking process", *Fuel*, Print ISSN: 0016-2361, Online ISSN: 1873-7153, 15 September 2024, Vol. 372, No. 1, pp. 132162, Published by Elsevier B.V., DOI: 10.1016/j.fuel.2024.132162, Available: https://www.sciencedirect.com/science/article/pii/S0016236124013103.

[10] Jo Plested, Musa Phiri and Tom Gedeon, "Deep transfer learning for image classification: a survey", *Artificial Intelligence Review*, Print ISSN: 0269-2821, Online ISSN: 1573-7462, 20 May 2022, Vol. 59, No. 1, pp. 100, Published by Springer, DOI: 10.1007/s10462-026-11491-z, Available: https://link.springer.com/article/10.1007/s10462-026-11491-z.

[11] Baoming Fu, "Variational Autoencoder-based High-dimensional Feature Extraction for Economic Analysis of Power Cost Data", *Informatica*, Print ISSN: 0868-4952, Online ISSN: 1854-8668, 7 March 2025, Vol. 49, No. 25, pp. 1–18, Published by Slovenian Society Informatika, DOI: 10.31449/inf.v49i25.8012, Available: https://informatica.vu.lt/journal/INFORMATICA/article/view/8012.

[12] Gangjin Wang, Yan Chen, You Zhu and Chi Xie, "Systemic risk prediction using machine learning: Does network connectedness help prediction?", *International Review of Financial Analysis*, Print ISSN: 1057-5219, Online ISSN: 1873-7153, 1 May 2024, Vol. 93, No. 1, pp. 103147, Published by Elsevier, DOI: 10.1016/j.irfa.2024.103147, Available: https://www.sciencedirect.com/science/article/abs/pii/S1057521924000796.

[13] Pan Tang, Tiantian Tang and Chennuo Lu, "Predicting systemic financial risk with interpretable machine learning", *The North American Journal of Economics and Finance*, Print ISSN: 1062-9408, Online ISSN: 1879-0054, 01 June 2024, Vol. 71, No. 1, pp. 102088, Published by Elsevier, DOI: 10.1016/j.najef.2024.102088, Available: https://www.sciencedirect.com/science/article/abs/pii/S1062940824000123.

[14] Ali Hadi Rabbani, Surayya Jamal, Muhammad Ali, Osama Ali, "Financial Risk Forecasting Using AI: Hybrid Econometric–Machine Learning Approaches", *The Critical Review of Social Sciences Studies*, Print ISSN: 2958-3215, Online ISSN: 2958-3223, 1 July 2025, Vol. 3, No. 4, pp. 795–809, Published by Critical Review Publishing, DOI: 10.59075/smv2vj25, Available: http://thecrsss.com/index.php/Journal/article/view/987.

[15] Yu Cheng, Zhen Xu, Yuan Chen, Yuhan Wang, Zhenghao Lin *et al.*, "A deep learning framework integrating CNN and BiLSTM for financial systemic risk analysis and prediction", in *Proceedings of the 2025 International Conference on Economic Management and Big Data Application*, 25-27 July 2025, Shenzhen, China, Print ISBN: 978-1-6654-5678-9, Online ISBN: 978-1-6654-5679-6, Vol. 1, No. 1, pp. 270–274, DOI: 10.1145/3770177.3770219, Available: https://dl.acm.org/doi/full/10.1145/3770177.3770219.

[16] Zeinab Srour, Jamil Hammoud and Mohamed Tarabay, "Forecasting Systemic Risk in the European Banking Industry: A Machine Learning Approach", *Journal of Risk and Financial Management*, Print ISSN: 1911-8074, Online ISSN: 1911-8082, 10 June 2025, Vol. 18, No. 6, pp. 335, Published by MDPI, DOI: 10.3390/jrfm18060335, Available: https://www.mdpi.com/1911-8074/18/6/335.

[17] Elric Donnelly and Liora Pendry, "Optimizing GAN-Based Data Augmentation for Predictive Financial Analytics", *Journal of Computer Technology and Software*, Print ISSN: 2998-2383, Online ISSN: 2998-2383, 1 August 2025, Vol. 4, No. 8, pp. 1–16, Published by Tech Science Press, DOI: 10.5281/zenodo.17074564, Available: https://ashpress.org/index.php/jcts/article/view/207.

[18] Jinhong Wu, Konstantinos Plataniotis, Lucy Liu, Ehsan Amjadian and Yuri Lawryshyn, "Interpretation for variational autoencoder used to generate financial synthetic tabular data", *Algorithms*, Print ISSN: 1999-4893, Online ISSN: 1999-4893, 20 February 2023, Vol. 16, No. 2, pp. 121, Published by MDPI, DOI: 10.3390/a16020121, Available: https://www.mdpi.com/1999-4893/16/2/121.

[19] Jie Hu, Li Shen and Gang Sun, "Squeeze-and-excitation networks", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 18-23 June 2018, Salt Lake City, UT, USA, Print ISBN: 978-1-5386-0458-5,

Online ISBN: 978-1-5386-0459-2, 18 June 2018, Vol. 1, No. 1, pp. 7132–7141, Published by IEEE Computer Society, DOI: 10.1109/CVPR.2018.00745, Available: https://ieeexplore.ieee.org/document/8578843.

[20] Ke Zhang, Miao Sun, Tony Xin Han, Xingfang Yuan, Liru Guo *et al*., "Residual networks of residual networks: Multilevel residual networks", *IEEE Transactions on Circuits and Systems for Video Technology*, Print ISSN: 1051-8215, Online ISSN: 1558-2205, 1 June 2017, Vol. 28, No. 6, pp. 1303–1314, Published by IEEE, DOI: TCSVT.2017.2654543, Available: https://ieeexplore.ieee.org/document/7820046.

[21] Hmidi Alaeddine and Malek Jihene, "Deep residual network in network", *Computational Intelligence and Neuroscience*, Print ISSN: 1687-5265, Online ISSN: 1687-5273, 23 February 2021, Vol. 2021, No. 1, Article No. 6659083, Published by Hindawi, DOI: 10.1155/2021/6659083, Available: https://www.hindawi.com/journals/cin/2021/6659083.

[22] Zhiwei Li, Guodong Long, Tianyi Zhou, Jing Jiang and Chengqi Zhang, "Personalized federated collaborative filtering: A variational autoencoder approach", in *Proceedings of the AAAI Conference on Artificial Intelligence*, 25 February 2025 – 4 March 2025, Philadelphia, Pennsylvania, Print ISSN: 2159-5399, Online ISSN: 2374-3468, Vol. 39, No. 17, pp. 18602–18610, Published by AAAI Press, DOI: 10.1609/aaai.v39i17.34047, Available: https://ojs.aaai.org/index.php/AAAI/article/view/34047.

[23] Zhangyu Mei, Xiao Bi, Dianguo Li, Wen Xia, Fan Yang *et al*., "DHHNN: a dynamic hypergraph hyperbolic neural network based on variational autoencoder for multimodal data integration and node classification", *Information Fusion*, Print ISSN: 1566-2535, Online ISSN: 1872-6305, 1 July 2025, Vol. 119, No. 1, Article No. 103016, Published by Elsevier B.V., DOI: 10.1016/j.inffus.2025.103016, Available: https://www.sciencedirect.com/science/article/pii/1566253525000892.

[24] Folasade Olubusola Isinkaye, Michael Olusoji Olusanya and Ayobami Andronicus Akinyelu, "A multi-class hybrid variational autoencoder and vision transformer model for enhanced plant disease identification", *Intelligent Systems with Applications*, Print ISSN: 2667-3053, Online ISSN: 2667-3053, 1 May 2025, Vol. 26, No. 1, Article No. 200490, Published by Elsevier B.V., DOI: 10.1016/j.iswa.2025.200490, Available: https://www.sciencedirect.com/science/article/pii/S266730532500016X.

[25] Asmaul Hosna, Ethel Merry, Jigmey Gyalmo, Zulfikar Alom, Zeyar Aung *et al*., "Transfer learning: a friendly introduction", *Journal of Big Data*, Print ISSN: 2196-1115, Online ISSN: 2196-1115, 1 December 2022, Vol. 9, No. 1, p. 102, Published by Springer, DOI: 10.1186/s40537-022-00652-w, Available: https://link.springer.com/article/10.1186/s40537-022-00652-w.

[26] Jaya Gupta, Sunil Pathak and Gireesh Kumar, "Deep learning (CNN) and transfer learning: a review", *Journal of Physics: Conference Series*, Print ISSN: 1742-6588, Online ISSN: 1742-6596, 1 September 2022, Vol. 2273, No. 1, Article No. 012029, Published by IOP Publishing, DOI: 10.1088/1742-6596/2273/1/012029, Available: https://iopscience.iop.org/article/10.1088/1742-6596/2273/1/012029.

[27] Kaggle, "Economic Resilience & Risk Modeling Dataset", *Kaggle Datasets*, 01 January 2025, Published by Kaggle Inc., DOI: 10.34740/kaggle.ds.222222.333, Available: https://www.kaggle.com/datasets/kickmuncher/economic-resilience-and-risk-modeling-dataset.

[28] Michael Greenacre, Patrick J. F. Groenen, Trevor Hastie, Alfonso Iodice D'Enza, Angelos Markos *et al*., "Principal component analysis", *Nature Reviews Methods Primers*, Online ISBN: 2662-8449, 22 December 2022, Vol. 2, p. 100, Published by Springer Nature, DOI: 10.1038/s43586-022-00184-w, Available: https://www.nature.com/articles/s43586-022-00184-w.

[29] Ahmed Mesai Belgacem, Mounir Hadef and Abdesslem Djerdir, "Fault diagnosis of photovoltaic arrays based on support vector machine and t-distributed stochastic neighbor embedding", in *Proceedings of the 5th International Conference on Electrical Engineering and Control Applications (ICEECA 2022) – Volume 2*, 15–17 November 2022, Khenchela, Algeria, Print ISBN: 978-981-97-4775-7, Online ISBN: 978-981-97-4776-4, Vol. 1224, No. 1, pp. 175–186, Published by Springer Singapore, DOI: 10.1007/978-981-97-4776-4_18, Available: https://link.springer.com/chapter/10.1007/978-981-97-4776-4_18.

[30] Tongyi Zhang, "Decision tree and k-nearest-neighbors (KNN)", in *An Introduction to Materials Informatics*, Singapore: Springer Nature, 27 February 2025, Online ISBN: 978-981-99-7992-9, ch. 7, pp. 117–140, Published by Springer Singapore, DOI: 10.1007/978-981-99-7992-9_5, Available: https://link.springer.com/chapter/10.1007/978-981-99-7992-9_5.

[31] Prashant Pawar, Shreyas Gonjari, Sushil Kshirsagar and Gajendra J. Chhajed, "A review of linear regression and support vector regression", *International Journal of Scientific Research in Engineering and Management*, Print ISSN: 2582-3930, Online ISSN: 2582-3930, 2 January 2025, Vol. 8, No. 12, pp. 1–9, Published by IJSRM Publisher, DOI: 10.55041/IJSREM40400, Available: https://ijsrem.com/download/a-review-of-linear-regression-and-support-vector-regression.

© 2025 by the author(s). Published by Annals of Emerging Technologies in Computing (AETiC), under the terms and conditions of the Creative Commons Attribution (CC BY) license which can be accessed at http://creativecommons.org/licenses/by/4.0.