Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 5, No. 7, 2026

Deep Representation Learning for Risk Prediction in Electronic Health Records Using Self-Supervised Methods

Anzhuo Xie[1*]
[1]Columbia UniversityNew York, USA
*Corresponding author: Anzhuo Xie; xieanzhuo@outlook.com

Abstract: This study proposes a self-supervised representation learning-based risk prediction model to address the challenges of label scarcity, structural complexity, and high heterogeneity in Electronic Health Records (EHR) data for risk assessment tasks. The model combines masked reconstruction and context prediction to automatically learn temporal dependencies and latent semantic structures in EHR data under unlabeled conditions, thereby generating robust and generalizable patient health representations. The overall architecture includes a feature embedding layer, a temporal encoding module, an attention-based context aggregation module, and a risk decoding layer, which together capture long-term dependencies and achieve hierarchical feature modeling from multidimensional medical information. Multiple experiments were conducted on the MIMIC-III dataset, and comparisons with LSTM, BiLSTM, Transformer, GRU, and GAT models show that the proposed method achieves significant improvements in Accuracy, F1-Score, Precision, and AUC, verifying the effectiveness of the self-supervised mechanism for EHR representation and risk assessment. In addition, sensitivity analyses of key hyperparameters such as learning rate, hidden layer dimension, and noise interference level demonstrate that the model maintains stable performance under various training conditions, reflecting strong structural robustness and noise-resistant feature extraction ability. Overall, this study achieves the goal of learning high-quality health representations from unlabeled EHR data and provides an efficient and scalable technical framework for intelligent medical risk prediction.

Keywords: Self-supervised representation learning; electronic health records; risk prediction; robust modeling

1. Introduction

In recent years, with the continuous advancement of medical informatization, Electronic Health Records (EHR) have become one of the most representative and valuable sources of medical data. EHR systems record multidimensional information generated during patient care, including demographic features, clinical activities, laboratory tests, prescriptions, imaging results, and vital signs[1]. These data are characterized by temporal continuity, heterogeneity, and high-dimensional sparsity, providing a rich foundation for disease risk assessment and personalized health management. Through deep mining of EHR data, healthcare systems are shifting from "experience-driven" to "data-driven" intelligent decision-making, enabling early warning, precise intervention, and dynamic monitoring. This transformation improves the efficiency of healthcare resource utilization and enhances patient outcomes. However, the complexity and incompleteness of EHR data make it difficult for traditional supervised learning models to capture underlying structures and semantic information, highlighting the urgent need for new representation learning paradigms to address large-scale yet sparsely labeled data[2].

In medical risk assessment, EHR data often suffer from missing labels, annotation bias, and uneven time spans. Since obtaining medical labels requires professional expertise and clinical validation, building large-scale labeled datasets is extremely costly. This limitation constrains the generalization and stability of supervised learning models. Moreover, EHR data exhibit significant heterogeneity[3]. Records from different hospitals, regions, or populations show statistical shifts and collection discrepancies, which further increase the difficulty of model transfer and domain adaptation. At the same time, the dynamic evolution of patient health states means that risk assessment depends not only on static features but also on modeling long-term dependencies and potential causal structures in time series. Therefore, how to automatically extract general and robust semantic representations under limited annotation becomes a key issue for improving risk prediction and model interpretability[4].

The rise of Self-Supervised Representation Learning (SSL) offers a promising solution to this challenge. Its core idea is to design auxiliary tasks that allow the model to learn structural and semantic information directly from data without external labels[5]. Compared with traditional feature engineering and supervised methods, self-supervised learning can better adapt to the multimodal and non-uniform characteristics of EHR data. It maintains strong generalization and transferability even in scenarios with limited labels. By learning latent representations, models can automatically capture hidden associations across different time slices and feature channels, providing a stable feature space for downstream risk assessment. This approach not only reduces dependence on manual annotations but also enables a unified representation framework across multiple centers and tasks, laying a foundation for cross-institutional risk modeling and intelligent decision-making[6].

In the field of health risk assessment, reliability, interpretability, and fairness are essential. EHR data often contain sensitive attributes such as age, gender, ethnicity, and socioeconomic status. Improper feature learning may lead to model bias and unfair decisions. The value of self-supervised learning in this context lies not only in enhancing feature expressiveness but also in improving robustness to imbalanced and noisy data through structural constraints and pretraining mechanisms. By pretraining on unlabeled data, models can form prior knowledge of health states and reduce prediction bias caused by sample imbalance or anomalies. Furthermore, self-supervised learning can be integrated with interpretability techniques to help clinical researchers understand the basis of model decisions, thereby improving usability and trustworthiness in real medical settings.

Overall, research on self-supervised representation learning based on Electronic Health Records represents a key shift in medical artificial intelligence from "task-driven" to "knowledge self-discovery." It provides a new paradigm for building intelligent, transparent, and fair health risk assessment systems. This approach effectively integrates temporal, structured, and unstructured information under limited supervision to achieve dynamic understanding of patients' multidimensional health states. Such research is of great significance for improving early disease screening, optimizing medical resource allocation, and promoting public health management. In the future, with the maturation of EHR data standardization and privacy protection technologies, self-supervised representation learning will become a core driving force for intelligent and precise healthcare, opening broader prospects for health risk prediction and decision support.

2. Related work

In recent years, risk assessment based on Electronic Health Records (EHR) has become one of the core directions in intelligent healthcare research[7]. Traditional approaches often rely on feature engineering and statistical modeling. They construct feature sets by selecting key clinical indicators, medication records, or laboratory results, and then apply models such as logistic regression or decision trees for classification and prediction. However, these methods depend heavily on manually designed features and fail to capture complex temporal dependencies and nonlinear relationships. With the introduction of deep learning, researchers began to adopt automatic feature extraction to learn high-dimensional latent representations directly from raw EHR data. Convolutional Neural Networks have been used to capture local diagnostic patterns, while Recurrent Neural Networks and their variants are effective in modeling temporal correlations within disease progression. This transition has significantly improved the accuracy and flexibility of risk prediction. Nevertheless, these methods usually rely on large labeled datasets, which are limited in real healthcare settings due to privacy concerns and the high cost of data annotation, resulting in poor model transferability and generalization[8].

Under limited annotation conditions, self-supervised learning has gradually emerged as an effective alternative. This approach designs pretraining tasks that enable the model to learn latent structures and semantic relationships from data without manual labels. Similar to its success in natural language processing and computer vision, self-supervised pretraining on EHR data can exploit internal dependencies within clinical sequences, such as diagnosis-medication co-occurrence, laboratory trend prediction, and masked visit-time reconstruction. Through these tasks, models can automatically learn intrinsic connections between disease evolution and health risks. Compared with traditional supervised feature learning, this approach achieves stable representation transfer across multiple institutions and data sources while mitigating distribution inconsistency caused by sampling bias among different hospitals. It provides a scalable feature foundation for downstream tasks such as risk assessment, disease prediction, and multitask joint modeling.

At the same time, researchers have explored multimodal EHR representation learning to enhance model understanding of complex medical scenarios. EHR data often include both structured information, such as laboratory values and prescription records, and unstructured information, such as clinical notes and imaging reports. Integrating these heterogeneous modalities within a unified framework has become a critical challenge. Self-supervised learning shows clear advantages in this regard. Through multi-view contrastive learning, cross-modal matching, and feature alignment mechanisms, models can capture associations among different modalities and enhance global semantic consistency. Representation learning frameworks based on these principles not only improve model generalization but also maintain stable predictions under missing or incomplete data conditions. Especially in tasks such as patient risk stratification and early disease detection, cross-modal self-supervised mechanisms can effectively integrate multidimensional medical evidence and provide more interpretable clinical insights[9].

In addition, recent research on model interpretability and fairness has been increasingly incorporated into self-supervised representation learning frameworks. Since medical data involve individual privacy and ethical considerations, model transparency and reliability are of great importance. By introducing attention constraints, contrastive consistency, or causal structural constraints, models can identify which features or temporal segments are most critical for risk assessment while learning latent representations. This helps reduce algorithmic bias and improves fairness across population groups. It also provides clinicians with traceable explanations for model decisions[10]. Overall, research in this field has evolved from feature engineering to representation learning and from supervised modeling to self-supervised and multimodal fusion frameworks. These developments are shaping a more intelligent, robust, and interpretable EHR-based risk assessment system, laying a solid technical foundation for precision medicine.

3. Method

Our framework aims to automatically extract temporal and hierarchical health risk feature representations from electronic health records (EHRs) through self-supervised representation learning, enabling high-quality risk assessment modeling. The overall approach comprises three core components: multi-source EHR feature modeling, self-supervised pre-training, and task-specific risk encoding. The model architecture is shown in

Figure 1. Specifically, the framework first integrates heterogeneous EHR data such as diagnoses, medications, and laboratory results into a unified temporal sequence representation, ensuring that temporal dependencies and cross-modal interactions are preserved. Then, through a self-supervised pre-training strategy combining masked reconstruction and contextual prediction, the model captures latent structural relationships and intrinsic dynamics within patient trajectories, forming robust and transferable feature embeddings. Finally, a task-specific risk encoding layer aggregates the learned temporal representations and maps them to individualized health risk scores, achieving an interpretable and data-efficient end-to-end modeling pipeline for clinical risk assessment.

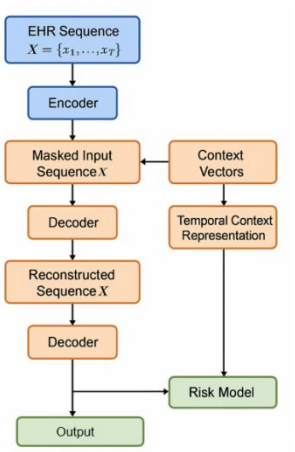

Figure 1. Overall model architecture

First, let the patient's diagnosis and treatment sequence within time interval $[0,T]$ be $X = \{x_1, x_2, ..., x_T\}$, where $x_t \in R^d$ represents the multidimensional medical features at time t. To capture temporal dependency and semantic consistency, the model introduces an embedding mapping function $f_\theta(\cdot)$ to project the original input into the latent representation space:

$$h_t = f_\theta(x_t) = RELU(W_e x_t + b_e) \tag{1}$$

Where W_e and b_e are bias terms, and d_h represents the hidden layer dimension. This representation serves as the basic feature vector for subsequent self-supervision tasks and risk modeling.

To achieve unsupervised representation of EHR temporal structure, the model adopts a self-supervised mechanism that combines mask reconstruction with temporal prediction. Specifically, a number of time slices in the sequence are randomly selected for masking to obtain the corrupted input sequence $\tilde{X}$, and the original signal is restored through the encoder-decoder framework:

$$\hat{X} = g_\phi(f_\theta(\tilde{X})) \tag{2}$$

Where $g_\phi(\cdot)$ represents the decoder, and the parameter set ϕ is optimized together with the encoder parameters θ. The self-supervised goal of the model is to minimize the reconstruction error:

$$L_{recon} = \frac{1}{T}\sum_{t=1}^{T} || x_t - \hat{x}_t ||_2^2 \tag{3}$$

This mechanism enables the model to learn temporal dependencies and global structural features in the absence of labels, enhancing the robustness and generalization ability of representation.

In time dimension modeling, in order to capture long-term dependencies, an attention-based context encoding strategy is introduced. The model calculates the context weight of each time step:

$$a_{t,i} = \frac{\exp(h_t^T W_a h_i)}{\sum_{j=1}^{T} \exp(h_t^T W_a h_j)} \tag{4}$$

Where W_a is a learnable matrix. The final temporal context vector is expressed as:

$$c_t = \sum_{i=1}^{T} a_{t,i} h_i \tag{5}$$

This mechanism can dynamically focus on key diagnostic and treatment segments, thereby identifying potential time-sensitive features in risk assessment, such as disease evolution trends and implicit pathological patterns.

In the risk modeling phase, the model aggregates the temporal features obtained from self-supervised learning into a global health representation z and implements risk encoding through nonlinear transformation. The global representation is obtained through weighted aggregation:

$$z = Aggregate(\{c_1, c_2, ..., c_T\}) = \frac{1}{T}\sum_{t=1}^{T} c_t \tag{6}$$

The model then outputs an individual risk index using a parameterized risk scoring function $r_\varphi(\cdot)$:

$$r = \sigma(w_r^T z + b_r) \tag{7}$$

Where $\sigma(\cdot)$ is a sigmoid function and w_r, b_r is a learnable parameter. Through this process, the model can form a stable and clinically meaningful health risk map under unlabeled conditions, providing a solid representation foundation for subsequent diagnostic assistance and personalized assessment.

4. Experimental Results

4.1 Dataset

This study uses the MIMIC-III (Medical Information Mart for Intensive Care III) dataset as the primary source for experiments and model analysis. The dataset consists of real electronic health records collected from intensive care units (ICUs). It contains multidimensional medical information for more than forty thousand patients, including demographic data, laboratory test results, prescriptions, vital signs, diagnosis codes, and dynamic events such as admissions and discharges. One notable feature of the MIMIC-III dataset is its time-series

structure. The physiological data of each patient are continuously recorded during hospitalization, reflecting the dynamic evolution of disease progression and risk states. This provides a rich temporal foundation for self-supervised representation learning.

The dataset exhibits both high dimensionality and heterogeneity. It includes quantitative physiological indicators as well as discrete medical event information, making it an important benchmark for studying the adaptability of self-supervised learning in medical representation tasks. To ensure data reliability and generalizability, strict preprocessing steps are commonly applied. These include missing value imputation, outlier detection, temporal alignment, and feature normalization, which together create a unified input space. The openness and standardization of MIMIC-III guarantee the reproducibility of models and provide a solid basis for transfer to other EHR platforms.

By applying self-supervised modeling to the continuous health records in MIMIC-III, the study can learn multi-level health representations under unlabeled conditions and uncover latent pathological patterns and risk signals. This representation learning, grounded in real clinical contexts, helps verify model robustness in complex healthcare environments and provides transferable knowledge for future cross-institutional and cross-task risk assessments. The extensive use of MIMIC-III and its high-quality multimodal features make it an essential experimental setting for applying self-supervised learning to electronic health record analysis.

4.2 Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table1. Comparative experimental results

Model	Accuracy	F1	Precision	Recall	AUC
LSTM[11]	0.861	0.854	0.849	0.860	0.887
BILSTM[12]	0.872	0.865	0.859	0.870	0.896
Transformer[13]	0.884	0.877	0.874	0.880	0.912
GRU[14]	0.868	0.860	0.857	0.864	0.891
GAT[15]	0.891	0.882	0.879	0.884	0.918
Ours	0.911	0.903	0.901	0.905	0.937

As shown in Table 1, different models exhibit significant variations in performance on risk assessment tasks based on electronic health records. Traditional temporal modeling methods such as LSTM and GRU have certain advantages in capturing changes in patients' health states. However, due to their limited ability to model long-term dependencies, these models face performance bottlenecks in complex disease progression representation. Although both models maintain high Recall values, indicating good sensitivity to potential risk events, their Precision is relatively low. This suggests that misclassification still occurs, and the models cannot fully capture the latent semantic associations among multidimensional medical features.

In contrast, the Transformer architecture performs better in modeling global dependencies through its self-attention mechanism. Its Accuracy and AUC are notably higher than those of recurrent neural network models. This indicates that attention-based sequence encoding can better capture potential causal relationships across different time intervals in EHR data, enabling stronger discriminative ability in risk prediction. Moreover, the representation power of the Transformer provides a better feature space foundation for downstream self-supervised tasks. However, its purely attention-based structure remains sensitive to noise and has limited generalization ability when handling high-dimensional heterogeneous medical data, making it difficult to adapt fully to inter-patient variability.

The GAT model introduces graph-structured relational modeling, which captures topological associations among medical events in the feature space. Experimental results show that the F1 and Precision values of GAT exceed those of the Transformer. This demonstrates that the graph-based feature propagation mechanism enhances the model's ability to represent interactions among patients' multidimensional features. Such a structure helps mitigate information loss in heterogeneous feature fusion for traditional sequential models, making the model more robust in identifying complex pathological patterns and multisource risk factors. However, GAT still faces limitations in capturing global temporal dependencies and needs further enhancement in modeling long time-span health records.

Overall, the proposed self-supervised representation learning model achieves the best performance across all metrics, particularly with Accuracy, F1, and AUC values of 0.911, 0.903, and 0.937, respectively. It demonstrates stronger stability and generalization capability. This advantage is mainly attributed to the model's ability to extract deep structural features of EHR data during pretraining through masking reconstruction and context prediction tasks. As a result, the model forms robust latent representations even with limited labeled data. These results confirm the effectiveness of the self-supervised mechanism in medical time-series data, showing that label-free representation learning can significantly enhance the model's ability to identify complex health risk patterns and provide new insights for intelligent and generalizable risk prediction.

This paper also presents a learning rate sensitivity experiment, the experimental results of which are shown in Figure 2. To further examine the stability and robustness of the proposed model, this experiment systematically explores how variations in the learning rate influence convergence dynamics, feature representation quality, and prediction consistency during training. By adjusting the learning rate across multiple magnitudes, the analysis aims to reveal the model's adaptability to optimization conditions and its ability to maintain performance under different parameter update scales. This setting provides an important perspective for understanding how the self-supervised representation learning framework responds to gradient variations in the context of EHR data modeling.

From the figure, it can be seen that the model's performance varies significantly under different learning rates, indicating that the learning rate plays an important role in the stability and convergence speed of the self-supervised representation learning process. When the learning rate is low (such as 1e-5), the model converges slowly but remains stable during training. The final Accuracy and F1-Score still stay at a relatively high level. As the learning rate increases to 1e-4, the model reaches its peak performance across all four metrics, suggesting that

this value provides a good balance between weight update efficiency and parameter stability.

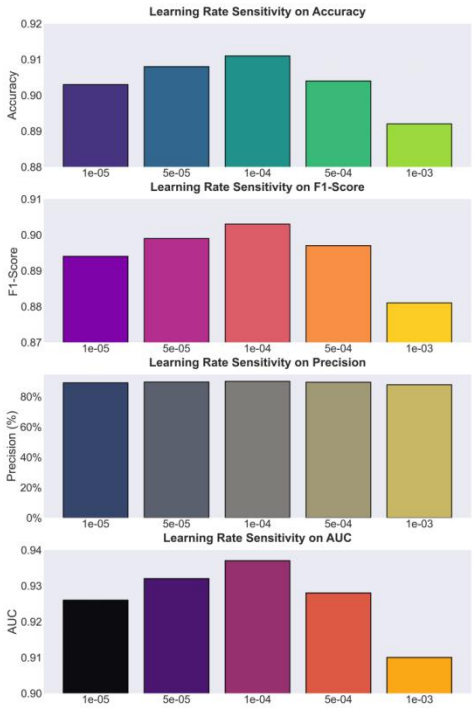

Figure 2. Learning rate sensitivity experiments

When the learning rate continues to increase to 5e-4 and above, the model's performance declines slightly, with noticeable decreases in Precision and AUC. This trend indicates that an excessively high learning rate causes rapid parameter updates, making it difficult for the model to capture the complex temporal dependencies and structural features in EHR data. Such behavior can lead to underfitting or gradient oscillations. In particular, for self-supervised reconstruction tasks, a large learning rate disrupts the consistency of the feature space, reducing the transferability and robustness of the learned representations.

The trends of different metrics show that Accuracy and F1-Score exhibit similar variation curves, while Precision is more sensitive to changes in the learning rate. This suggests that in risk assessment tasks, the learning rate not only affects overall classification performance but also directly determines the accuracy of identifying high-risk samples. The fluctuation of the AUC metric further supports this observation. When the learning rate is too high, the model's ability to distinguish risk samples drops significantly, indicating that its decision boundaries in the feature space become unstable, making it difficult to effectively separate risk and normal samples.

Overall, the experimental results demonstrate that the self-supervised representation learning framework is highly sensitive to the learning rate setting. A properly chosen learning rate can significantly improve model performance in EHR-based risk prediction. The optimal learning rate, around 1e-4, enables the model to achieve the best balance between reconstruction and risk encoding, forming a more robust health representation structure. This finding suggests that in self-supervised EHR modeling, optimizing hyperparameters not only affects convergence speed but also directly determines the effectiveness of latent feature extraction and the precision of subsequent risk identification.

This paper further presents a hidden layer dimension sensitivity experiment, the experimental results of which are shown in Figure 3.

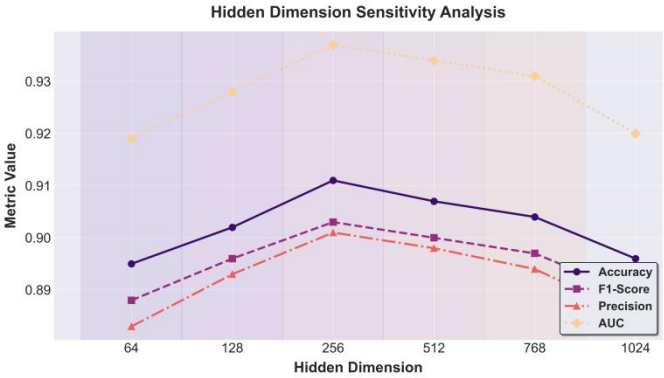

Figure 3. Hidden layer dimension sensitivity experiment

From the figure, it can be observed that the model shows a clear nonlinear trend under different hidden layer dimensions, indicating that the hidden layer dimension plays a critical role in the capability of self-supervised representation learning. When the dimension is small (such as 64 or 128), both Accuracy and F1-Score remain low, suggesting that the feature space capacity is insufficient to effectively capture the complex temporal dependencies and multimodal semantic information in electronic health records. As the dimension increases to 256, the model reaches its peak performance. This shows that the latent space at this scale achieves an optimal balance between representational richness and parameter stability, allowing the model to achieve higher discriminative power in risk assessment tasks.

When the hidden layer dimension further increases to 512 or above, the performance metrics begin to decline gradually, with noticeable decreases in Precision and AUC. This trend indicates that excessively high dimensions may lead to feature redundancy and noise amplification. As a result, the model tends to overfit the reconstruction task during self-supervised pretraining, weakening the discriminative ability of its learned features. The feature diffusion effect in high-dimensional space reduces the model's focus on key clinical variables, which negatively impacts the accuracy of downstream risk assessments. Moreover, higher dimensions may increase optimization difficulty, causing gradient instability and unnecessary computational overhead during training.

Overall, the experimental results demonstrate that a proper hidden layer dimension is essential for constructing high-quality self-supervised health representations. The optimal dimension, around 256, enables the model to efficiently learn dynamic features of patient health states in the latent space while maintaining good generalization and stability. This finding further confirms the rationality of the proposed structural design. By balancing feature capacity and parameter complexity, the model achieves robust modeling and efficient risk identification for EHR data, providing a strong foundation for future transfer and extension across various healthcare tasks.

This paper also presents a noise interference level sensitivity experiment, the experimental results of which are shown in Figure 4.

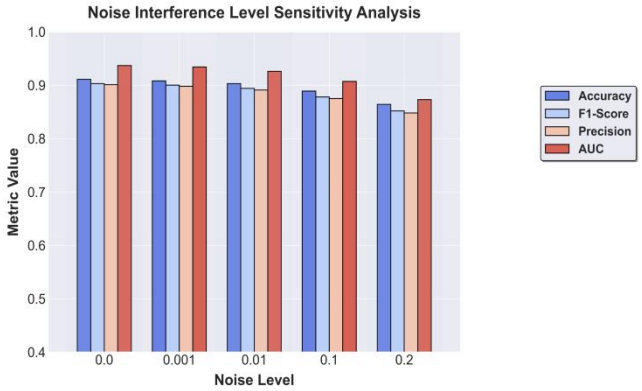

Figure 4. Noise interference level sensitivity experiment

From the figure, it can be seen that as the noise level increases, all performance metrics show a slight downward trend but remain within a high range. This indicates that the proposed self-supervised representation learning framework demonstrates strong robustness to input noise. When the noise level is low (between 0.0 and 0.01), Accuracy and F1-Score remain almost stable. This suggests that the model can effectively capture the main semantic information in electronic health records and suppress minor random disturbances during feature extraction. It also reflects that through masked reconstruction and context prediction during pretraining, the model has learned to adaptively repair local anomalies and missing information.

When the noise level rises to 0.1, Precision and AUC begin to decline more noticeably. This shows that noise disturbances gradually disrupt the consistency of the latent feature space, causing some samples to drift near the risk boundaries. This phenomenon is common in medical data, as noise in EHRs often originates from abnormal entries, measurement errors, or recording biases. The proposed model still maintains a relatively high AUC level, indicating that it retains strong discriminative power in risk identification. This can be attributed to the global feature alignment mechanism introduced during self-supervised training, which helps the model preserve stable discriminative features even under high-noise conditions.

As the noise ratio further increases to 0.2, the model performance declines more significantly, especially in F1-Score and Precision. This trend indicates that under high-noise conditions, the model's ability to balance recall and precision becomes more challenging. Some low-confidence samples may be misclassified, affecting overall predictive performance. Nevertheless, the model does not experience catastrophic degradation, which shows that its adaptive mechanism under multi-noise environments remains effective. Comparing performance across different noise levels suggests that the robustness of the model mainly comes from the compression of redundant features and the reinforcement of key variables during multitask self-supervised training.

Overall, the experimental results confirm the noise-resistant stability of the proposed method in EHR-based risk assessment tasks. Proper self-supervised feature modeling not only improves learning efficiency under low-noise conditions but also maintains robust representational capacity under strong noise interference. This ability to suppress noise disturbance makes the model more reliable and generalizable in real-world medical applications, providing solid technical support for analyzing complex, multisource, and unstructured EHR data.

5. Conclusion

This study constructs a self-supervised representation learning framework based on electronic health records to systematically explore the feasibility and advantages of unsupervised feature modeling in health risk assessment. The results show that the proposed method can effectively learn multidimensional patient health features under unlabeled conditions, capturing temporal dependencies and cross-modal semantic associations. It achieves stable and accurate risk prediction in complex medical environments. Compared with traditional supervised models, the proposed framework demonstrates clear advantages in feature expressiveness, generalization performance, and noise robustness, providing new methodological support for the sustainable development of medical artificial intelligence.

At the methodological level, the self-supervised learning mechanism proposed in this study uses masked reconstruction and context prediction tasks to enable the model to discover latent structural patterns and temporal dependencies in raw EHR data. This feature learning approach enhances the model's representational capacity and provides transferable health representations for downstream tasks such as early disease detection, hospitalization risk assessment, and treatment pathway optimization. In addition, the systematic analysis of key hyperparameters, including learning rate, hidden layer dimension, and noise interference level, further validates the model's robustness and stability. The results indicate that self-supervised learning can maintain strong performance under various conditions, offering practical evidence for its application in real-world medical AI systems.

At the application level, this research has significant implications for intelligent healthcare, clinical decision support, and health management. Self-supervised representation learning can overcome the limitation of scarce labeled data, enabling medical models to generalize well across diverse institutions and populations. This contributes to building interpretable and scalable medical risk assessment systems and promotes the transition of healthcare from experience-driven to data-driven decision-making. More importantly, the proposed framework can assist physicians in early risk identification and personalized diagnosis and treatment. It improves clinical efficiency, reduces misdiagnosis rates, and supports the optimal allocation of healthcare resources through data-driven insights.

Looking ahead, as electronic health records continue to grow and data-sharing mechanisms improve, self-supervised representation learning will further expand its research and application value in intelligent risk assessment. Future work can be deepened in several directions. First, integrating multimodal medical information such as imaging, genomics, and clinical text can achieve cross-modal joint representation. Second, incorporating causal inference and uncertainty estimation mechanisms can enhance model interpretability and

decision reliability. Third, exploring privacy-preserving and federated learning approaches can ensure patient data security while enabling collaborative model optimization across institutions. Overall, this study provides a new paradigm and technical pathway for intelligent medical risk assessment and offers important insights for building secure, trustworthy, and efficient medical AI systems in the future.

References

[1] Yao H R, Cao N, Russell K, et al. Self-supervised representation learning on electronic health records with graph kernel infomax[J]. ACM Transactions on Computing for Healthcare, 2024, 5(2): 1-28.

[2] Kumar Y, Ilin A, Salo H, et al. Self-supervised forecasting in electronic health records with attention-free models[J]. IEEE Transactions on Artificial Intelligence, 2024, 5(8): 3926-3938.

[3] Wang X, Luo J, Wang J, et al. Hierarchical pretraining on multimodal electronic health records[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing. 2023, 2023: 2839.

[4] AlSaad R, Malluhi Q, Abd-Alrazaq A, et al. Temporal self-attention for risk prediction from electronic health records using non-stationary kernel approximation[J]. Artificial Intelligence in Medicine, 2024, 149: 102802.

[5] Wornow M, Xu Y, Thapa R, et al. The shaky foundations of large language models and foundation models for electronic health records[J]. npj digital medicine, 2023, 6(1): 135

[6] Nayebi Kerdabadi M, Hadizadeh Moghaddam A, Liu B, et al. Contrastive learning of temporal distinctiveness for survival analysis in electronic health records[C]//Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. 2023: 1897-1906.

[7] Zhong Y, Cui S, Wang J, et al. Meddiffusion: Boosting health risk prediction via diffusion-based data augmentation[C]//Proceedings of the 2024 SIAM International Conference on Data Mining (SDM). Society for Industrial and Applied Mathematics, 2024: 499-507.

[8] Zhang Z, Cui H, Xu R, et al. Tacco: Task-guided co-clustering of clinical concepts and patient visits for disease subtyping based on ehr data[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2024: 6324-6334.

[9] Heilbroner S P, Carter C, Vidmar D M, et al. A self-supervised framework for laboratory data imputation in electronic health records[J]. Communications Medicine, 2025, 5(1): 251.

[10] Lee S, Yin C, Zhang P. Stable clinical risk prediction against distribution shift in electronic health records[J]. Patterns, 2023, 4(9).

[11] Liu L J, Ortiz-Soriano V, Neyra J A, et al. KIT-LSTM: knowledge-guided time-aware LSTM for continuous clinical risk prediction[C]//2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2022: 1086-1091.

[12] Pang H, Zhou L, Dong Y, et al. Electronic health records-based data-driven diabetes knowledge unveiling and risk prognosis[J]. arXiv preprint arXiv:2412.03961, 2024.

[13] Wang, Q., X. Zhang and X. Wang, "Multimodal integration of physiological signals, clinical data and medical imaging for ICU outcome prediction", Journal of Computer Technology and Software, vol. 4, no. 8, 2025.

[14] Sottile P D, Albers D, DeWitt P E, et al. Real-time electronic health record mortality prediction during the COVID-19 pandemic: a prospective cohort study[J]. Journal of the American Medical Informatics Association, 2021, 28(11): 2354-2365.

[15] Piya F L, Gupta M, Beheshti R. Healthgat: Node classifications in electronic health records using graph attention networks[C]//2024 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). IEEE, 2024: 132-141.