

Citation: Lavanya M, Mangayarkarasi S. A deep auto encoder based security mechanism for protecting sensitive data using AI based risk assessment. *Journal of Harbin Institute of Technology (New Series)*. DOI: 10.11916/j.issn.1005-9113.2024018.

A Deep Auto Encoder Based Security Mechanism for Protecting Sensitive Data Using AI Based Risk Assessment

Lavanya M and Mangayarkarasi S

(Department of Computer Science, Vels Institute of Science Technology & Advanced Studies (VISTAS), Pallavaram, Chennai 600117, Tamilnadu, India)

Abstract: Big data has ushered in an era of unprecedented access to vast amounts of new, unstructured data, particularly in the realm of sensitive information. This data presents unique opportunities for enhancing risk alerting systems, but also poses challenges in terms of extraction and analysis due to its diverse file formats. This paper proposes the utilization of a DAE-based (Deep Auto Encoders) model for projecting risk associated with financial data. The research delves into the development of an indicator assessing the degree to which organizations successfully avoid displaying bias in handling financial information. Results from simulation experiments demonstrate the superior performance of the DAE algorithm, showcasing fewer false positives, improved overall detection rates, and a noteworthy 9% reduction in failure jitter. The optimized DAE algorithm achieves an accuracy of 99%, surpassing existing methods, thereby presenting a robust solution for sensitive data risk projection.

Keywords: data mining; sensitive data; DAE

CLC number: TP39.F832.59

Document code: A

Article ID: 1005-9113(2024)00-0000-09

0 Introduction

The systemic financial risk is commonly used to refer to potential threats that could impact the entire monetary system. One prevalent type of systemic risk in the financial sector is exemplified by the recent financial crisis^[1], which serves as a typical illustration of this risk. Since the 17th century, there has been an ongoing increase in the frequency of financial crises worldwide, along with an increase in their severity.

Although there has been some recent improvement in the international financial situation, the global financial market is still adjusting to new conditions and recovering from the effects of the crisis. Both the probability of exogenous financial dangers, such as those presented by the globalization of the economy, and the severity of their impact, are increasing. One reason for this is the increased interconnectedness of many economies due to globalization^[2].

In recent years, due to increasing scientific and

technological capabilities, the country has been at the forefront of constant innovation and the development of new forms of financial systems. This position has allowed China to remain a leader in constant innovation. For instance, alternative payment systems^[3] have begun to replace more conventional ones in the realm of digital finance in recent years.

Additionally, significant progress has been made in digital insurance, intelligent investment, and online financing. The financial system is becoming more vulnerable to disruptions originating within the nation. Because of the nature of the Internet, threats can quickly spread across various sectors and regions, eventually becoming financial risks, posing a risk to businesses.

Anticipating potential dangers to financial well-being is extremely challenging. The ineffectiveness of traditional approaches to financial risk largely stems from the absence of essential factors that are both efficient and timely in their application, making early warning technology unable to provide reliable early warnings.

Both academics and businessmen agree that the characteristics of models ultimately determine their implementation in real life. The traditional early warning technology for financial risk relies on information and factors generated from conventional statistical data at the factor level, enabling more accurate predictions^[4]. There is no specific stance either in favor of or against the disclosure of prospective financial risks.

The advent of big data has brought an abundance of new, unstructured data that can be leveraged to enhance financial risk alerting systems. This data comes in various file formats and can be mined in multiple ways. In recent years, significant progress has been made in artificial intelligence, particularly in vision, natural language comprehension, and cognitive perception^[5]. These advancements enable the mining of data, leading to the identification of key financial risk warning factors that are both efficient and timely.

This research not only provides applicable algorithms for predicting financial risk but also discusses the widespread use of artificial intelligence in applications involving the mining of image and text data.

Tools such as scanners, Optical Character Recognition (OCR) programs, and Natural Language Processing (NLP) software can extract valuable information from scanned documents and images, thus enabling early warning systems^[6]. Optical character recognition allows the extraction of risk assessment information from non-standard sources^[7-8]. Utilizing data from remote sensing enables dynamic predictions regarding population density and urban development rates^[9].

The use of speech-to-text recognition technology can enhance the interactive experience and improve safety levels in financial applications^[10]. Technology employing NLP and machine learning can extract real-time insights regarding financial entities, correlations among financial events, and factors indicating economic uncertainty from textual data sources such as news, public opinion, and forums^[11].

Images and texts represent novel data sources, posing challenges due to their diverse origins, types, large volumes, and high frequency. Unlike traditional data collected mainly by governments and institutions, image and text big data are multisource and heterogeneous^[12]. The absence of standardized collection formats for unstructured data poses

significant challenges for AI information gathering and preprocessing technologies.

Conventional methods of data collection relying on paper media can only handle a limited amount of information due to financial constraints. The processing speed of unstructured data is becoming increasingly crucial as large datasets containing images and text can take seconds or longer to process. Given the variables involved, applying unstructured big data to provide early warnings of potential financial risks is extremely challenging. It is essential for risk warning purposes to accurately and effectively extract usable information from high-frequency, heterogeneous data originating from various sources^[13]. Our research introduces a novel approach by integrating advanced deep learning techniques, specifically Deep Auto Encoders (DAE), into the realm of financial risk assessment within the IoT ecosystem. This novel integration not only improves the accuracy and efficiency of risk quantification but also addresses challenges such as false detections, workload reduction, and enhanced precision without introducing unnecessary complexity. By pioneering this innovative methodology, our work contributes significantly to advancing risk management strategies in complex and interconnected digital environments.

1 Related Works

The ability to make decisions is one of the most crucial skills for management in any company. Each passing day adds a new layer of complexity to the business world.

Making decisions is becoming an increasingly challenging and unpredictable process due to this complexity. Businesses of all sizes and types, along with various other organizations, are turning to decision support systems for assistance when faced with difficult decisions. Awareness of these systems is growing every day. Given the multitude of approaches to overcoming challenges, it is crucial to utilize a wide range of strategies.

Yazdani et al.^[14] conducted research on Quality Function Deployment (QFD) to discover the most effective approach for the agricultural supply chain. Pancreatic islet transplantation has been suggested as a potential treatment option by Scalia et al.^[15], who presented a multi-factor evaluation approach for the procedure. Additionally, a fuzzy system was

established to assist with decision-making in situations where there is uncertainty regarding a medical diagnosis.

Cobreria-Paniagua et al.^[16] proposed the use of artificial emotions to enable an autonomous emotional decision-making system, aiming to increase the system's level of autonomy in investment decisions. This is recommended to enhance the system's autonomy level. It has also been recommended that the goals of an adaptive stock index trading decision system should include predicting the future direction of stock index prices, capitalizing on trading opportunities, and minimizing losses to the greatest extent possible. These are just a few of the recommendations made.

The Financial Expert System (FES) established by Weng et al.^[17] analyzed the counts and sentiment ratings of news articles to forecast short-term stock prices. Another system designed to assist users in financial trading predicts stock trading patterns using a combination of support vector machines and portfolio selection theory. The goal of this system is to help users make more informed trading decisions by analyzing previously collected data.

Most of these earlier studies reflected the business viewpoint without considering the investor's perspective simultaneously. It is necessary to consider the role of investors, who are among the most important players in the financial industry, to arrive at a fair conclusion.

Businesses must consider various components and maintain awareness of their surroundings to identify potential dangers. Research on creating systems that can anticipate and alert businesses to impending dangers has been ongoing for some time.

Early Warning Systems (EWS) are tools that can help individuals and organizations mitigate risks by forecasting unusual circumstances. They assist in quantifying risks based on current circumstances and potential future threats. One of the many applications of EWS is early detection and reporting of financial issues, risks, and opportunities that could impact financial statements. This application is just one of many, as EWSs offer possibilities for protection and problem mitigation^[18].

The EWS used in the financial business is shaped by combining several models. Typically, logistic and probit regression models are utilized in conventional EWS practices. A model can be created by incorporating various indicators to account for country-

specific variations. Methods such as Multiple Criteria Decision Making (MCDM), which involve decision-making based on multiple criteria, have also been employed^[19]. Additionally, there has been consideration given to the idea of developing an EWS using simulated neural networks to some extent.

Kim et al.^[20] utilized an Artificial Neural Network (ANN) as an economic predictor while applying EWS models to the Korean economy. This approach allowed the researchers to better analyze the data.

Yang et al.^[21] identified prospective banking risks, however, the studies focused primarily on factors indicating the status of businesses, neglecting the perspectives of investors. This enterprise-centered evaluation approach may skew the findings, highlighting the need for a solution that incorporates investor perspectives.

Hirra et al.^[22] focused on classification techniques using patch-based deep learning modeling. Qadri et al.^[23] introduced SVseg, a stacked sparse autoencoder-based patch classification modeling approach for vertebrae segmentation. This work contributes to the field of medical imaging and computer-aided diagnosis by developing advanced algorithms for accurate segmentation tasks. The study by Ref. [24] presented a CT-based automatic spine segmentation method using patch-based deep learning. Ahmad et al.^[25] delved into facial expression recognition using lightweight deep learning modeling. This research explored the practical applications of deep learning in computer vision tasks, particularly in recognizing and interpreting facial expressions, which has implications in human-computer interaction.

Previous research proposed a novel risk detection strategy by integrating investor perspectives. This approach involved using measures that represented investors' viewpoints alongside organizational risk assessments. By combining business risk levels with investor opinion levels, a more comprehensive risk assessment could be achieved. Following this, an analysis of business patterns or circumstances was conducted to determine the risk level. Investment plans were then provided based on the chosen risk level preferred by the investor.

2 Methodology

Studies, aimed at identifying potential crises, have

the primary purpose of locating potential warning signs by making extensive use of the data affiliated with the company. Disclosure data and financial statements provide an objective description of the company. As a result, these two types of information are generally employed in the process of locating risk signals.

It is possible to reconstruct the previous activities and occurrences of businesses in relation to credit with the assistance of these data. They can provide us with assistance in evaluating the health of companies and determining which ones require the most attention from our end. In other words, they can help us identify the businesses that require the most focus from us.

It may be possible to determine whether a business can be relied upon using data such as financial statements. However, there may be risks that are not disclosed, such as significant events or accounting manipulation. Information that is not related to finances, such as the feelings of investors, can serve as a leading indicator of potential problems for these businesses.

Opinion mining is utilized by businesses of this phenomenon to ascertain information that had been concealed in the past or to gauge the feeling of consumers. An enterprise-specific risk warning indicator is created by combining these two different kinds of information into a single metric.

Since the risk indicator defining characteristic is the possibility of a credit incident, that aspect ought to be the primary focus of the indicator attention. The calculated indicator is another tool that can be used in the process of determining the level of risk that the business is subjected to. The final step in the process of different business processes guarantees that the organisation will have the ability to recognise risk signals in real time.

As shown in Fig. 1, this research considers not one but two distinct perspectives when analysing a crisis. These perspectives are as follows: (1) the aspect of corporate data known as opinion mining, which can objectively describe the nature of the business; and (2) the points of view of shareholders. If a thorough risk evaluation is performed, the result may be a more precise estimation of the extent to which the company is vulnerable to risk.

2.1 Preprocessing

The unprocessed data contains gaps that need to be filled with appropriate information before conducting any further research. Columns with a

substantial amount of missing data are removed, along with specific samples lacking integrity to handle missing units. This process eliminates columns with significant data gaps. Each segment of information corresponds to different characteristics, and attributes with substantial data gaps are excluded from the analysis. Rows represent individual samples, and in the final analysis, samples with significant missing data are penalized and considered poor predictors. Rows themselves symbolize the samples.

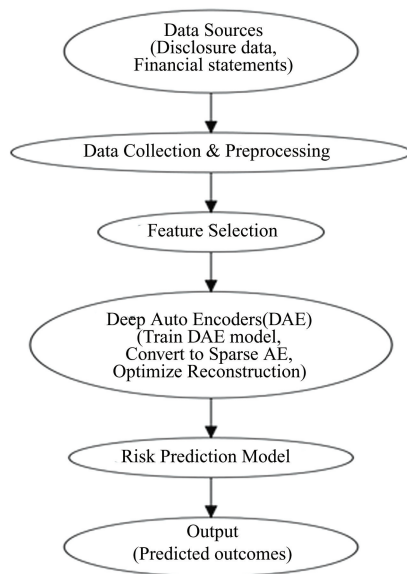


Fig.1 Risk detection in financial transaction

Vertical and lateral data cleansing are processes carried out to ensure that data can be utilized in subsequent evaluations. When compared to the entire data collection, the amount of material removed from consideration is insignificant. Upon completing the data cleansing process, we did not observe a significant reduction in the sample size we had been working with.

A sufficient amount of sampling was conducted to ensure that the final product would not be significantly altered due to these relatively small alterations. This was achieved by sampling an adequate quantity of the final product. During our data collection process, a few categories were left incomplete, resulting in no corresponding numbers assigned to them. We have concluded that the proportion of unused occurrences is low enough to justify removing the NULL values from the dataset. This is the conclusion we have reached.

Data deletion is carried out based on the percentage of rows and columns containing NULL

values. This is done to minimize the risk of inadvertently eliminating any underlying correlations within the data. However, before selecting a machine learning model to use, we need to determine which features are crucial for building our final model.

2.2 Feature Selection

It has come to our attention that the information we have collected contains a significant amount of redundant details, and it is imperative to eliminate these redundancies. The outcome of the prediction will not be influenced by the overwhelming majority of these variables, and some may even complicate the analysis.

Filtering is especially crucial for features that represent details that are irrelevant on their own. If you seek a more convincing and effective analysis beyond simply examining each feature's relevance individually, mathematical analysis should be considered. This approach considers a greater number of variables, leading to a more comprehensive analysis.

The selection of features for data analysis directly impacts the accuracy of the model. If the filtering process does not significantly affect model accuracy, it is optimal to use a smaller number of features when choosing between dozens or hundreds of features.

To address this need, the Pearson correlation coefficient was developed. This figure helps determine which factors play the most significant role in the investigation's outcome with greater precision. Utilizing Pearson correlation coefficient technique can reduce the number of variables necessary for the model. The Pearson correlation coefficient is an indicator that shows how closely connected the two variables are to one another. To determine the Pearson correlation coefficient, merely take the correlation and divide it by the standard deviation, as shown in Eq.(1).

$$p_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y} \tag{1}$$

On the other hand, it would be more natural to approach specific instances in the following manner, as shown in Eq.(2).

$$r = \frac{N \sum_i x_i y_i - \sum_i x_i \sum_i y_i}{\sqrt{N \sum_i (x_i)^2 - \sum_i (x_i)^2} \sqrt{N \sum_i (y_i)^2 - \sum_i (y_i)^2}} \tag{2}$$

The value of the variable can range from -1 to 1, with larger absolute values indicating a higher level of

correlation. We can adjust the threshold value to control the number of features most related to each other, allowing us to identify closely associated characteristics. This filtering process has led to increased productivity and efficiency. Each dataset has been cleaned, standardized, and formatted to ensure accuracy in subsequent analysis models. Preprocessing the data is always the first and crucial step to ensure trustworthy and unbiased outcomes in further investigations.

2.3 Deep Auto Encoders

It is the function of neural networks such as autoencoders to store the data that is input into the network so that it can be reconstituted as output data later. The autoencoder must first acquire the knowledge necessary to determine how to isolate the primary characteristics from the input using the train set $\{x(1), x(2), \dots, x(n)\}$, where $x(i) \in \mathbf{R}^d$

Fig. 2 illustrates the basic autoencoder that has one input layer, one disguised layer, and one output layer. The autoencoder model starts by encoding the single input $x(i)$ to the hidden layer $y(x(i))$, which is then decoded as the output layer $z(x(i))$, which concludes that the autoencoder model was successful.

$$y(x) = f(W_1 x + b) \tag{3}$$

$$z(x) = g(W_2 x + c) \tag{4}$$

where W_1 means weight matrix, b is encode vector, W_2 is decode matrix, and c is decoding vector.

The logistic sigmoid function is expressed as below in Eq.(5).

$$f(x) = 1/(1 + \exp(-x)) \tag{5}$$

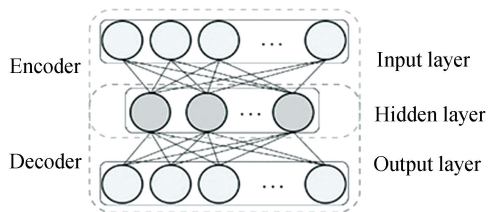


Fig. 2 Autoencoder processes

The autoencoder model requires an input layer, denoted by x , as well as an encoding function, denoted by f . This combination ultimately leads to the generation of an approximated output layer (y). The autoencoder model makes use of a decoder function (g) to reconstruct the original input layer (x), which ultimately leads to the generation of the output layer (z). Scaling with the loss function $LH(x,z)$ is what determines the reconstruction error. Finding the optimal settings for the reconstruction parameters is

accomplished by decreasing this function as $L(X, Z)$ and then finding the optimal settings for the reconstruction parameters.

$$\theta = \arg_{\theta} \min L(X, Z) = \arg_{\theta} \min 0.5 \sum_{i=1}^N \|x^{(i)} - z(x^{(i)})\|^2 \quad (6)$$

The size of the autoencoder model hidden layer, which is typically either equal to or larger than the size of the model output layer, is a significant concern in this field of research. The structure of the components that go into making up the model typically takes this problem into consideration. The autoencoder model was converted into a sparse autoencoder by utilizing a nonlinear autoencoder with a hidden layer that is one unit larger than the input layer and the sparsity constraint technique. The sparsity restriction method was utilized to successfully complete this task. We made use of a sparsity prerequisite so that we could receive a representation that is sparse and reduce the amount of error that was introduced during the reconstruction process.

$$S = L(X, Z) + \gamma \sum_{i=1}^{H_D} KL(\rho \|\rho^j) \quad (7)$$

$$\hat{\rho}_j = \frac{1}{N} \sum_{i=1}^N y_j(x^{(i)}) \quad (8)$$

where H_D means hidden units, γ is weight, ρ is sparsity parameter.

2.4 Score Based DAE

The DAE and its parameter configuration for congestion prediction can be evaluated in comparison to two cutting-edge deep learning neural network-based models that serve as benchmarks. This allows the effectiveness of the DAE and its parameter configuration for predicting congestion to be determined. To ensure that we can carry out the most exhaustive set of comparisons and evaluations feasible, we have settled on the strategy of making our projections based on three separate time horizons. The study evaluates the efficacy of traffic congestion predictions by utilizing the Mean Absolute Error (MAE) and the weighted Mean Squared Error (wMSE). The actual level of traffic congestion (ct_{ij} , bt_{ij} , and wt_{ij}), the anticipated level of traffic congestion, and the punishment weight at time t in the grid are all indicated by the coordinates (i, j) in the 2D-matrix representation of the highway transportation network. Additionally, the actual level of congestion is indicated by the coordinates (i, j) in the grid. The information can be viewed using this matrix as a

display. If uneven distributions of congestion intensity are a problem, one approach to the problem is to heuristically define wt_{ij} in such a way as to increase the penalty for inaccurate projections of non-smooth congestion levels. The breadth (W) and height (H) of the grid are represented by the abbreviations that correspond to them, respectively.

$$MAE = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H \|c t_{ij} - \hat{c} t_{ij}\| \quad (9)$$

$$wMSE = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H w t_{ij} \times (c t_{ij} - \hat{c} t_{ij})^2 \quad (10)$$

Algorithm 1: Financial risk prediction using DAE.

Input: Financial and disclosure statements.

Output: Risk prediction.

- 1) Pre-processing the input datasets.
- 2) Feature extract and select the required classes.
- 3) Classify the features using DAE.

$$\theta = \arg_{\theta} \min L(X, Z)$$
- 4) Score based evaluation using MAE (Eq.(9)) and wMSE (Eq.(10)).
- 5) Obtain the Predicted outcomes with minimal errors.

3 Results and Discussions

Data is selected from the KDD CUP99 (13.3%) dataset and fed into the Deep Auto Encoder (DAE) algorithm. The algorithm is implemented in Python, allowing us to check the detection rate and error detection rate. After completing the training phase, we will analyze the results obtained from the test dataset used during the network training process. Twelve separate experiments were conducted using data from the test set, and the detection rate and error detection rate of the DAE algorithm are compared with those of the existing algorithm.

The DDoS detection module is thoroughly tested using both strategies, and the experiment outcomes are presented in Figs.3 – 6. One desirable quality of the DAE algorithm is its low probability of detection for both false positives and false negatives. The research compares the results and explores the causes of the risks measured using these two distinct approaches.

The research has leveraged various IoT-powered applications to assess, monitor, and investigate potential financial threats. Moving forward, we plan to combine multiple algorithms for comparison, validate them with data, and develop a comprehensive

algorithm to identify and calculate financial risk indices. This algorithm will be used to investigate both the safety and threats to financial institutions.

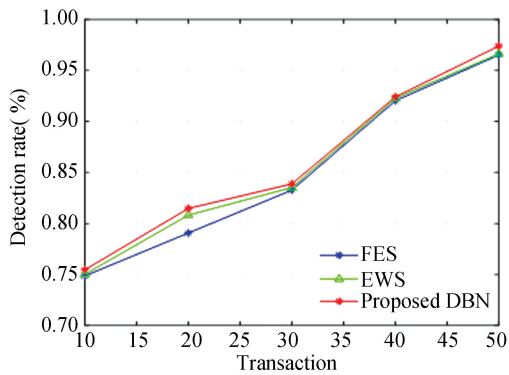


Fig.3 Detection rate

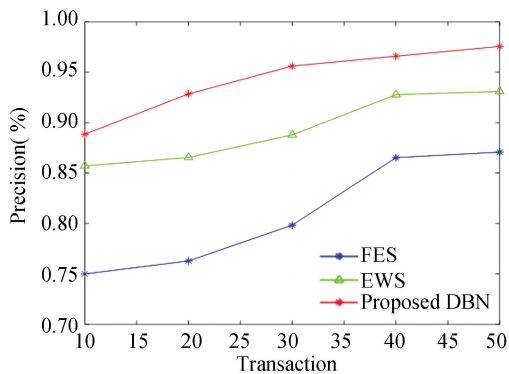


Fig.4 Precision

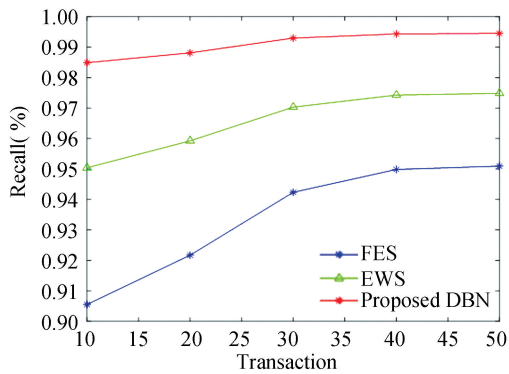


Fig.5 Recall

The network configuration comprises a total of 64 nodes, and we conducted a simulation of the Improved Deep Auto Encoder (IDAE) algorithm, comparing its results with those generated by the fundamental DAE algorithm. The increase in the number of connections in the network has led to a rise in the objective function's value. However, the rate of growth experienced by the IDAE algorithm is significantly slower than that of the DAE algorithm. The optimal reaction generated by the IDAE algorithm is of superior quality compared to that of the basic

DAE algorithm.

The IDAE algorithm demonstrates a significant reduction in delay jitter and can provide deterministic real-time assurances to time-sensitive networks. Comparing its results with other algorithms, we find that the IDAE algorithm has reduced delay jitter by 7.1% compared to the DAE algorithm under stable parameters.

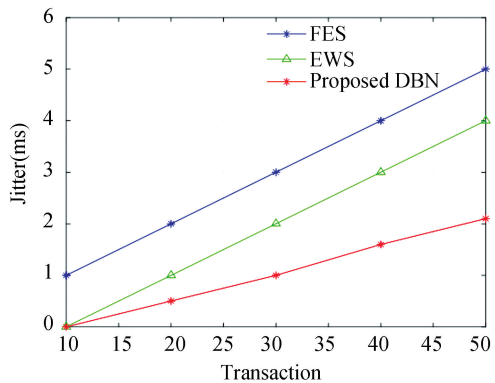


Fig. 6 Jitter

Improving financial risk assessment and management is a critical aspect of modern finance, especially in the context of evolving technologies and interconnected global markets. In recent years, researchers and practitioners have increasingly turned to advanced computational techniques, such as deep learning algorithms, to enhance the accuracy and efficiency of financial risk detection and analysis. One such promising algorithm is the IDAE algorithm, which builds upon the capabilities of the DAE algorithm to better handle large-scale financial data and improve risk quantification.

To evaluate the effectiveness of the IDAE algorithm in financial risk assessment, we implemented various configurations and conducted 75 risk measurements across networks of different sizes, ranging from 5 to 55 links. This comprehensive approach allowed us to assess the algorithm's performance across a spectrum of network complexities and sizes.

One notable observation from our experiments is the relationship between network size and algorithm performance. As expected, the total time required to execute both DAE and IDAE algorithms increases linearly with the growth in network size. However, we observed an interesting trend regarding running times; while the IDAE algorithm took longer than DAE algorithm for smaller-scale network configurations, it exhibited shorter running times for

slightly larger network topologies. This trend suggests that IDAE algorithm's scalability and efficiency improve as the complexity of the network increases, highlighting its suitability for large-scale risk evaluation tasks.

IDAE algorithm's superior performance in handling larger network topologies can be attributed to its incremental improvements over DAE algorithm. By incorporating enhanced computational techniques and optimized network configurations, IDAE algorithm demonstrates improved risk quantification capabilities and faster processing times for complex financial data sets.

Moreover, our findings underscore IDAE algorithm's advantages in terms of potential risk capture, maximum processing speed, and overall stability. These qualities are crucial for financial institutions and risk managers who must analyze vast amounts of data accurately and swiftly to make informed decisions and mitigate potential risks effectively.

The training and validation loss curves provide insights into the learning dynamics of the machine learning model. The graph below (Fig.7) depicts the training and validation loss trends over epochs during model training.

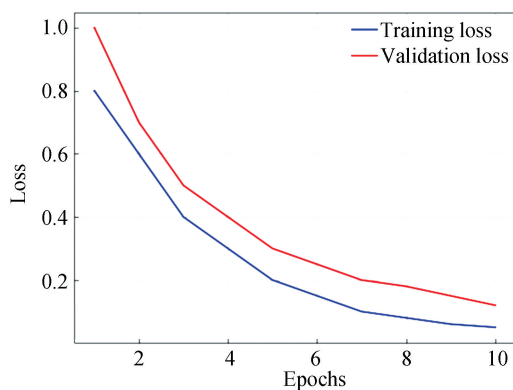


Fig.7 Training and validation loss

In addition to performance metrics, such as running times and scalability, we also evaluated the IDAE algorithm's effectiveness in capturing and quantifying various types of financial risks. The algorithm's ability to analyze diverse risk factors and provide meaningful insights is essential for comprehensive risk management strategies.

Furthermore, our experiments highlight the importance of continuous advancements in computational techniques for financial risk assessment. As financial markets evolve and become more

interconnected, traditional risk assessment methods may prove inadequate in capturing emerging risks and market dynamics. Advanced algorithms like the IDAE algorithm offer promising solutions by leveraging machine learning and deep learning principles to uncover hidden patterns and correlations within complex financial data.

It is important to note that while IDAE algorithm demonstrates significant improvements over existing approaches, there are still challenges and limitations to consider. For instance, the computational resources required for running the algorithm on large-scale datasets can be substantial, necessitating efficient hardware infrastructure and optimization techniques.

Moreover, as with any algorithmic approach, the IDAE algorithm's effectiveness relies on the quality and relevance of the input data. Data preprocessing, cleaning, and feature selection remain critical steps in ensuring the algorithm's accuracy and reliability in risk assessment tasks.

4 Conclusions

This article explores into the changes occurring in financial risk within the IoT ecosystem. It employs various algorithms to assess and quantify these risks, identifies potentially malicious nodes in the risk landscape, analyzes the contributing factors to these risks, and ultimately assigns an overall risk ranking. All of these critical aspects are thoroughly discussed and detailed in this article. The results from the experiments suggest that, under specific circumstances, IDAE algorithm outperforms DAE algorithm in measuring financial risk, demonstrating a lower rate of false detection. This improvement not only reduces the workload required but also enhances the precision of risk assessment, all without introducing unnecessary complexity into the program. Integrating this performance-enhancing algorithm into consecutive measurements can lead to optimal results in risk assessment tasks. Moreover, the ability to pinpoint malicious nodes within the risk landscape adds an extra layer of security and risk management capability. By accurately identifying and addressing potential threats, organizations can proactively protect their financial systems and data assets. The comprehensive analysis provided in this article serves as a valuable resource for researchers, practitioners, and decision-makers in the finance and IoT industries.

It underscores the importance of leveraging advanced algorithms like DAE algorithm to navigate and mitigate evolving financial risks in interconnected systems.

References

- [1] Li Y. Security and risk analysis of financial industry based on the internet of things. *Wireless Communications and Mobile Computing*, 2022, 2022: Article ID 6343468. DOI: 10.1155/2022/6343468.
- [2] Mhlanga D. Industry 4.0 in finance: The impact of Artificial Intelligence (AI) on digital financial inclusion. *International Journal of Financial Studies*, 2020, 8(3): 45. DOI: 10.3390/ijfs8030045.
- [3] DuHadway S, Carnovale S, Hazen B. Understanding risk management for intentional supply chain disruptions: Risk detection, risk mitigation, and risk recovery. *Annals of Operations Research*, 2019, 283:179–198. DOI: 10.1007/s10479–017–2452–0.
- [4] Motoc M M. A Proposal for a bankruptcy risk detection model-adaptation of the Taffler model. *Ovidius University Annals, Economic Sciences Series*, 2021,21(2):406–412.
- [5] Khan A B, Devi S, Devi K. An enhanced AES-GCM based security protocol for securing the IoT communication. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, 23(4): 711–719. DOI: 10.17586/2226–1494–2023–23–4–711–719.
- [6] Wang D X, Lin J B, Cui P, et al. A semi-supervised graph attentive network for financial fraud detection. 2019 IEEE International Conference on Data Mining (ICDM). Piscataway: IEEE. 2019:598–607. DOI: 10.1109/ICDM.2019.00070.
- [7] Papadakis S, Garefalakis A, Lemonakis C, et al. *Machine Learning Applications for Accounting Disclosure and Fraud Detection*. Philadelphia: IGI Global. 2020. DOI: 10.4018/978–1–7998–4805–9.
- [8] Xie H, Shi Y. A big data technique for internet financial risk control. *Mobile Information Systems*, 2022, 2022(1): Article ID 9549868. DOI: 10.1155/2022/9549868.
- [9] Rehman H U. Financial risks classification early warning analysis of data mining technology. *Journal of Global Humanities and Social Sciences*, 2022, 3(3): 57–60. DOI:10.47852/bonviewGHSS2022030305.
- [10] Zhang S. Research on enterprise financial risk prediction method based on regression analysis. *SHS Web of Conferences*. 2023, 154: 02017. DOI: 10.1051/shsconf/202315402017.
- [11] Toma F M, Cepoi C O, Kubinski M N, et al. Gazing through the bubble: An experimental investigation into financial risk-taking using eye-tracking. *Financial Innovation*, 2023, 9: 28. DOI: 10.1186/s40854–022–00444–4.
- [12] Peng S, Yang S, Yao J. Improving value-at-risk prediction under model uncertainty. *Journal of Financial Econometrics*, 2023, 21(1): 228–259. DOI: 10.1093/jffinec/nbaa022.
- [13] Xia Y, Xu T, Wei M X, et al. Predicting chain's manufacturing SME credit risk in supply chain finance based on machine learning methods. *Sustainability*, 2023, 15(2): 1087. DOI:10.3390/su15021087.
- [14] Yazdani M, Chatterjee P, Zavadskas E K, et al. Integrated QFD-MCDM framework for green supplier selection. *Journal of Cleaner Production*, 2017, 142(Part 4): 3728–3740. DOI:10.1016/j.jclepro.2016.10.095.
- [15] Scalia G L, Aiello G, Rastellini C, et al. Multi-criteria decision making support system for pancreatic islet transplantation. *Expert Systems with Applications*, 2011, 38(4): 3091–3097. DOI:10.1016/j.eswa.2010.08.101.
- [16] Cabrera-Paniagua D, Cubillos C, Vicari R, et al. Decision-making system for stock exchange market using artificial emotions. *Expert Systems with Applications*, 2015, 42(20): 7070–7083. DOI:10.1016/j.eswa.2015.05.004.
- [17] Weng B, Lu L, Wang X, et al. Predicting short-term stock prices using ensemble methods and online data sources. *Expert Systems with Applications*, 2018, 112: 258–273. DOI: 10.1016/j.eswa.2018.06.016.
- [18] Koyuncugil A S, Ozgulbas N. Financial early warning system model and data mining application for risk detection. *Expert systems with Applications*, 2012, 39(6): 6238–6253. DOI: 10.1016/j.eswa.2011.12.021.
- [19] Kou G, Peng Y, Wang G. Evaluation of clustering algorithms for financial risk analysis using MCDM methods. *Information Sciences*, 2014, 275: 1–12. DOI: 10.1016/j.ins.2014.02.137.
- [20] Kim T Y, Oh K J, Sohn I, et al. Usefulness of artificial neural networks for early warning system of economic crisis. *Expert Systems with Applications*, 2004, 26(4): 583–590. DOI: 10.1016/j.eswa.2003.12.009.
- [21] Yang B, Li L X, Ji H, et al. An early warning system for loan risk assessment using artificial neural networks. *Knowledge-Based Systems*, 2001, 14(5–6): 303–306. DOI: 10.1016/S0950–7051(01)00110–1.
- [22] Hirra I, Ahmad M, Hussain A, et al. Breast cancer classification from histopathological images using patch-based deep learning modeling. *IEEE Access*, 2021, 9: 24273–24287. DOI: 10.1109/ACCESS.2021.3056516.
- [23] Qadri S F, Shen L, Ahmad M, et al. SVseg: stacked sparse autoencoder-based patch classification modeling for vertebrae segmentation. *Mathematics*, 2022, 10: 796. DOI: 10.3390/math10050796.
- [24] Qadri S F, Lin H X, Shen L L, et al. CT-based automatic spine segmentation using patch-based deep learning. *International Journal of Intelligent Systems*, 2023, 2023(1): Article ID 2345835. DOI: 10.1155/2023/2345835.
- [25] Ahmad M, Saira Alfandi O, Khattak A M, et al. Facial expression recognition using lightweight deep learning modeling. *Mathematical Biosciences and Engineering*, 2023, 20(5): 8208–8225. DOI: 10.3934/mbe.2023357.