IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 November 2024, accepted 8 January 2025, date of publication 13 January 2025, date of current version 23 January 2025.
Digital Object Identifier 10.1109/ACCESS.2025.3528373

RESEARCH ARTICLE

A Novel Hybrid Model for Credit Risk Assessment of Supply Chain Finance Based on Topological Data Analysis and Graph Neural Network

KOSAR FARAJPOUR MOJDEHI, BABAK AMIRI, AND AMIRALI HADDADI
School of Industrial Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran
Corresponding author: Babak Amiri (babakamiri@iust.ac.ir)

ABSTRACT Supply Chain Finance (SCF) in the energy sector has emerged as a critical area of focus due to the need for sustainable and efficient financial solutions to manage the complex interactions between various stakeholders, including suppliers, financial institutions, and energy companies. This study proposes a novel hybrid Topological Data Analysis (TDA) and Graph Neural Network (GNN) to optimize credit risk assessment in SCF. By leveraging BallMapper (BM) topological data analysis model and network-based features, the proposed model provides deeper insights into credit risk factors, enhancing the accuracy and dependability of credit risk evaluation for SMEs. Results demonstrate that the proposed BallMapper-Graph Neural Network (BM-GNN) model achieves higher accuracy and F1-scores, outperforming traditional machine learning approaches. Notably, incorporating network-based features alongside financial ratios yields the most favorable results in credit risk assessment. The SHapley Additive exPlanations (SHAP) model highlights the pivotal role of certain features in predicting bankruptcy, offering valuable insights for risk mitigation strategies. These results contribute to the growing body of evidence supporting the efficacy of TDA and GNN in financial applications, particularly in credit risk evaluation for SMEs in supply chain finance. Using network-based models opens up new avenues for improving accuracy and reliability in risk assessment, ultimately empowering financial institutions and stakeholders to make more informed decisions.

INDEX TERMS Topological data analysis, graph neural network, credit risk, BallMapper, supply chain finance.

I. INTRODUCTION

In the intricate landscape of the energy sector, SCF emerges as a crucial tool, particularly for SMEs, to navigate financial challenges and optimize cash flow [1], [2]. SMEs in the energy sector, including industries like oil, gas, petrochemicals, and power plants, face numerous challenges such as high capital requirements, complex regulatory compliance, market volatility, rapid technological advancements, and fierce competition from large enterprises [3]. Accessing sufficient financing is particularly difficult, as banks and financial institutions often view SMEs as high-risk [4]. Effective credit risk assessment is crucial for these SMEs as it helps improve their access to finance by providing a clearer picture of their financial health to lenders and investors [5], [6]. It aids in managing financial risks, enhancing business credibility, guiding strategic decisions, ensuring regulatory compliance, and attracting investment. These benefits are essential for the growth, sustainability, and competitive edge of SMEs in this capital-intensive and highly regulated industry. Traditional credit risk assessment methods, such as Decision Trees (DT), Support Vector Machines (SVM), and Logistic Regression (LR), primarily focus on financial ratios and linear relationships [7]. However, these models often fall short in capturing the complex, non-linear, and topological characteristics inherent in financial data, particularly within the networked structure of the energy sector's supply chain. To address these limitations, our research proposes a novel hybrid model that integrates BallMapper (BM), a topological data analysis method, with GNNs, which

The associate editor coordinating the review of this manuscript and approving it for publication was Aysegul Ucar.

excel in modeling dependencies and interactions within graphs.

The motivation behind combining BM and GNN lies in their complementary strengths. BM is adept at identifying patterns and clusters within high-dimensional data, revealing topological insights that traditional models overlook [8]. On the other hand, GNNs leverage relational information within graphs to model the dependencies and interactions among entities, capturing the contagion risk that arises from financial distress spreading through the supply chain network. By integrating these two methodologies, we aim to develop a more comprehensive and accurate credit risk assessment model, specifically tailored for the complexities of SCF in the energy sector.

Our research introduces the BM-GNN model, which offers several unique contributions and highlights its critical role in various segments of the energy industry:

A. ENHANCED FEATURE SELECTION

Traditional credit risk assessment methods often rely on linear analysis of relationships between financial variables and fail to model the complexities of multidimensional and non-linear data. BallMapper, as a topological analysis method, can uncover hidden structures and non-linear relationships in data. By selecting features analogous to Altman's Z-Score, it enhances the precision of predicting companies' financial statuses.

Hypothesis 1: The initial features (Basic Feature Set) are insufficient for accurate credit risk prediction, and selecting features using BallMapper will significantly improve model performance.

B. NETWORK-BASED CREDIT RISK ASSESSMENT

In supply chains, financial crises in one company can propagate to others, creating systemic risks. GNN models, leveraging network information, are capable of modeling these dependencies and complex interactions. Combining BallMapper with GNN enables the extraction of network-based features for a more accurate representation of risk propagation within the energy supply chain.

Hypothesis 2: Network-based features extracted using BallMapper better model the interdependencies and systemic risks in the energy supply chain compared to non-network features.

C. IMPROVED PREDICTIVE PERFORMANCE

Previous studies, such as [9], have demonstrated the success of GNN models in credit risk prediction. However, integrating topological methods like BallMapper with GNN can further strengthen the model's ability to identify non-linear relationships and enhance predictive accuracy. This research assumes that BallMapper-extracted features outperform correlation-based features in predicting bankruptcy.

Hypothesis 3: BallMapper-extracted features outperform correlation-based features in capturing non-linear relationships and improving predictive accuracy.

D. EXPLAINABILITY

One of the key challenges in complex machine learning models is transparency and explainability of predictions. In this study, the SHAP model is used to analyze the importance of features in credit risk prediction. Combining features selected by BallMapper with network information extracted by GNN provides more robust and transparent justifications for predictions, improving decision-making processes for managers and investors.

Hypothesis 4: Combining BallMapper-selected features and network-based features leads to the best overall performance, offering superior accuracy and explainability in credit risk prediction.

The remaining sections are structured as follows: Section II presents an in-depth exploration of relevant literature concerning credit risk prediction in SCF. We aim to provide an extensive understanding of the prevalent trends in related fields and to identify gaps in existing research that we intend to address. Section III outlines the development process of our proposed model, BallMapper GNN. We delve into three key stages of the model: feature extraction using BallMapper, graph transformation via clustering ideology, and node classification by Graph Neural Network. Moving on to Section IV, we offer a detailed account of the study conducted, which is divided into four phases. Initially, we provide a holistic overview of the data under examination. Subsequently, we elaborate on the implementation process of the model. Following this, we subject the implemented models to evaluative scrutiny, systematically comparing the obtained results. Ultimately, we summarize the key findings of our study and analyze its practical implications for management, in Section V.

II. LITERATURE REVIEW

Supply chain finance is a critical component in optimizing financial flows within supply chains, attracting significant attention from academia and industry [10]. It involves the use of various instruments and techniques such as reverse factoring and dynamic discounting to reduce inefficiencies in financial flows, resulting in substantial cost savings for networks [11]. Supply chain finance is essential for ensuring sustainable financial flows within industries, particularly in the face of challenges like financial crises that can impact supply chains [12]. The concept of supply chain finance is closely linked to supply chain risk management (SCRM), which is crucial for understanding and mitigating risks associated with financial flows and transactions within supply chains [13], [14]. In the realm of supply chain finance, Artificial Intelligence (AI) networks have been identified as a significant strategy for ensuring sustainable financial flows within the food and drink industry [12]. AI technologies are utilized to implement advanced solutions that enhance financial processes and improve the overall efficiency of supply chain finance operations. Additionally, blockchain technology has been recognized as a valuable tool for enhancing various functions within supply chains, including finance operations [15]. Blockchain technology offers opportunities to enhance transparency, traceability, and security in financial transactions within supply chains, thereby contributing to more efficient supply chain finance practices [16]. Efforts to redesign supply chains for the Circular Economy have highlighted challenges that impact supply chain finance [17]. These challenges underscore the need for innovative financial strategies to support sustainable practices within circular supply chains, ranging from product recovery processes to closed-loop supply chain network design. Furthermore, the application of game theory in supply chain finance has been explored to analyse the value of advance payment financing in reducing carbon emissions and improving production efficiency within supply chains [18]. Supply chain finance is intricately connected to supply chain coordination, where coordinating financial processes among supply chain members is crucial for maximizing the profitability of the entire supply chain [19]. Trade credit, when combined with effective coordination contracts, can significantly enhance the profitability of both individual members and the overall supply chain [20]. Moreover, the adoption of AI in supply chain management has the potential to revolutionize financial processes, although critical success factors influencing the adoption of AI in supply chain finance require further exploration [21]. The resilience of supply chains is another critical aspect that intersects with supply chain finance, especially in the face of increasing complexities and global uncertainties [22]. Proactive management practices are essential for enhancing supply chain resilience, particularly in the context of global sourcing strategies that introduce additional complexities to financial flows within supply chains. Additionally, the dynamic nature of supply chain networks necessitates the development of dynamic recovery policies to address disruptions and mitigate the ripple effect on supply chain economic performance [23].

Credit risk in supply chain finance is a crucial aspect that necessitates comprehensive understanding and management to ensure the smooth flow of financial transactions within the supply chain. Supply chain finance (SCF) has emerged as an innovative solution aimed at optimizing financial flows within supply chains [10]. However, the presence of credit risk can significantly impact the effectiveness of SCF. Credit risk refers to the potential that a borrower or counterparty will fail to meet its financial obligations, leading to financial losses for the lender or supplier. In the context of supply chain risk management (SCRM), financial risk encompasses various factors associated with financial flows and transactions within the supply chain [14]. This includes risks related to payment delays, defaults, insolvencies, and fluctuations in currency exchange rates. Managing financial risks effectively is crucial for maintaining the stability and efficiency of supply chain operations. One of the key challenges in addressing credit risk in supply chain finance is the lack of visibility and transparency in the financial flows within the supply chain network [24]. Without clear visibility of the financial transactions and relationships between different entities in the supply chain, it becomes difficult to assess and mitigate credit risk effectively. The emergence of longer and geographically dispersed supply chains has further exacerbated this challenge, exposing companies to a wider range of financial risks [24]. To mitigate credit risk in supply chain finance, it is essential to adopt proactive risk management strategies that involve identifying, assessing, and monitoring potential risks associated with financial transactions [25]. Artificial intelligence (AI) has been identified as a promising technology for enhancing supply chain risk management practices, including the identification and mitigation of credit risk [25]. By leveraging AI algorithms and predictive analytics, companies can improve their ability to assess creditworthiness, detect anomalies in financial transactions, and make data-driven decisions to mitigate credit risk effectively. Furthermore, building resilient supply chains through supplier flexibility and reliability assessment can also help mitigate credit risk in supply chain finance [26]. By developing strong relationships with reliable suppliers and fostering flexibility in the supply chain network, companies can reduce their exposure to credit risk and ensure continuity in financial transactions. Supplier evaluation mechanisms based on Bayesian belief networks can also aid in analysing and quantifying credit risk during supplier selection processes [27]. Offringa et al. [52] categorized risks in supply chain financial business into credit risk, market risk, operational risk, and systemic risk, emphasizing credit risk as a primary concern for risk management in supply chain finance [28]. It was highlighted that commercial banks play a pivotal role in evaluating and mitigating credit risks within the supply chain. Various models and methodologies have been proposed to tackle credit risk in supply chain finance. For example, Chen and Yano (2018) suggested streamlining ownership transfer procedures to reduce credit risk, emphasizing operational efficiency and transparency [29]. Xia et al. (2020) introduced a credit risk evaluation index system tailored to supply chain finance business characteristics, aiding in identifying and managing credit risks effectively [30]. Technological advancements like blockchain and the Internet of Things (IoT) have influenced credit risk management in supply chain finance [31]. Research has explored the integration of blockchain and fuzzy neural networks for supply chain financial risk assessment, showcasing the potential of advanced technologies in enhancing risk evaluation processes [32]. Additionally, the measurement of supply chain finance credit risk based on IoT has been investigated, highlighting the role of technological innovations in refining credit risk assessment methodologies within supply chain finance [33]. Predictive modelling has emerged as a valuable tool for assessing credit risk in supply chain finance. Models such as IG-GA-SVM and K-Means-SVM have been developed to forecast and manage credit risks effectively [34]. External factors, such as the COVID-19 pandemic, have also impacted credit risk in supply chain finance. Studies have shown increased credit risk contagion in supply chain finance during the pandemic, emphasizing the need for adaptive risk management strategies during times of economic uncertainty.

Based on the examination of the relevant literature, three primary viewpoints arise regarding potential pathways for enhancing predictive accuracies in credit risk assessment models. Firstly, the acknowledgment of the significant advantages of incorporating TDA models with other algorithms has demonstrated promising potential, especially through the application of graph representation learning models. Secondly, addressing scenarios involving implicit relations is highlighted. While Prior research on graph-based methodologies often rely on explicit relations, the challenge lies in scenarios where explicit relations are difficult to collect or acquire. There is a noticeable gap in research focusing on improving forecasts without explicit external relations. Thirdly, in contemporary forecasting methodologies, the integration of various analytical tools emerges as a fundamental aspect. These integrated models demonstrate exceptional effectiveness in improving forecasting accuracy, particularly when they are adaptable to diverse datasets and sources of information. This emphasis on integration highlights a significant transition toward more resilient and adaptable forecasting approaches within modern research frameworks.

In the context of credit risk assessment in supply chain finance, recent advancements in machine learning and graph-based methods provide novel pathways for enhancing predictive accuracy and risk evaluation processes. Traditional algorithms such as LR, DT, Multi Layer Perceptrons (MLP), and SVM have been extensively utilized due to their interpretability, robustness, and ability to handle various types of data. Each of these methods has its strengths and weaknesses, making them suitable for different contexts within credit risk evaluation.

Logistic regression remains one of the most widely used statistical methods for credit risk assessment. Its advantages include strong interpretability and the ability to provide probability scores for default risk, which are crucial for decision making in financial institutions. For instance, Zhou et al. demonstrated that an LR model achieved a prediction accuracy of 81.25% in assessing the credit risk of listed companies in China, highlighting its effectiveness in practical applications [35]. Additionally, Pederzoli and Thomas noted that while LR models do not require strict assumptions about the distribution of feature variables, they are limited by their capacity to include only a small number of predictors, which can affect their predictive power [36]. This limitation is echoed in the work of Wei and Hasan, who emphasized the utility of LR in commercial banking for assessing borrower default risk [37]. Decision Trees (DT) are another popular choice for credit risk assessment due to their intuitive structure and ease of interpretation. Guo et al. constructed a credit risk assessment model using DT alongside LR, demonstrating its effectiveness in evaluating companies with and without credit records [38]. The ability of DTs to visualize decision paths makes them particularly useful for stakeholders who require transparency in the decision making process. Multi-layer perceptron, a type of artificial neural network, have also been applied in credit risk assessment. Assef et al. compared MLP with LR and found that MLP could effectively analyze complex relationships in data, particularly in distinguishing between default and non-default borrowers [37]. However, MLPs often require larger datasets and more computational resources, which can be a drawback in certain contexts. Support Vector Machines (SVM) are recognized for their robustness in high dimensional spaces and their ability to handle nonlinear relationships. Nehrebecka's study compared SVM with LR in predicting default risk, finding that SVM models often outperformed LR in terms of predictive accuracy, particularly when dealing with complex datasets [39]. This aligns with the findings of Huang, who noted the effectiveness of SVM in evaluating credit risk for listed companies [40].

The integration of graph representation learning, such as Graph Convolutional Neural Networks (GCNNs) and GraphSAGE models, has proven effective in addressing complex relational structures within data. For instance, these methods have been applied to cybersecurity attack detection in cloud computing networks, showcasing their potential in identifying patterns and anomalies within dynamic and interconnected systems [41]. Similarly, adaptive learning techniques, such as transfer adaptation methods for feature expansion in multi-label deep neural networks, highlight the importance of dynamic feature extraction and the ability to adapt to evolving data landscapes [42]. These approaches underline the necessity of flexible and scalable methodologies in domains characterized by diverse and dynamic datasets. Moreover, collaborative frameworks for risk assessment, such as those developed for dynamic federations of cloud networks, emphasize the value of integrating distributed information sources to enhance decision-making under uncertainty [43]. This collaborative risk assessment aligns closely with the challenges of credit risk management in supply chain finance, where diverse financial flows and dependencies necessitate robust and coordinated analytical tools. By leveraging hybrid models that combine TDA methods like BallMapper with GNNs, our proposed BM-GNN framework builds on these advancements to address the complexities of credit risk prediction. This model bridges the gap between implicit relational extraction and the dynamic nature of financial networks, offering a resilient and adaptable solution tailored to the needs of SCF credit risk assessment.

We propose a novel hybrid model, BM-GNN, which uniquely combines Topological Data Analysis and Graph Neural Networks for a more nuanced and accurate credit risk assessment of SMEs. Unlike traditional approaches, our model incorporates network-based features that capture systemic interdependencies and topological insights, significantly enhancing predictive accuracy. The integration of SHAP explanations further ensures model interpretability, facilitating more informed decision-making for stakeholders.

III. PROPOSED MODEL

In this study, we propose the BM-GNN model for credit risk assessment, which integrates topological data analysis BallMapper with GNN. This model is designed to leverage the strengths of both approaches: BallMapper excels in identifying patterns and clusters in high-dimensional data, while GNN effectively analyzes complex relationships and dependencies within graph structures.

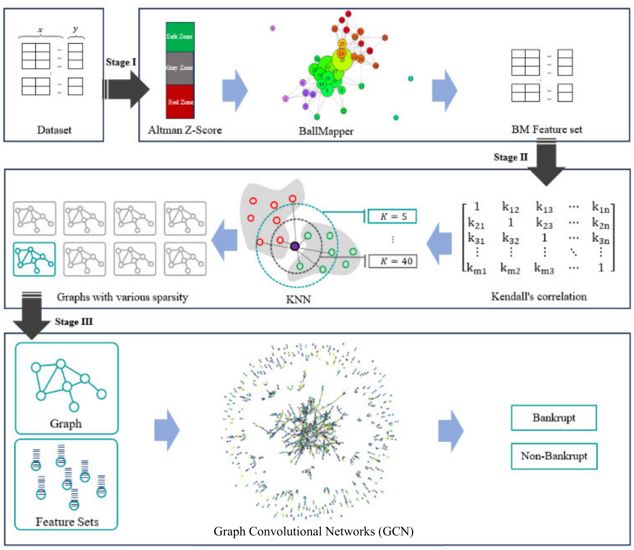


FIGURE 1. Proposed BM-GNN algorithm.

As shown in Fig. 1, the BM-GNN model is structured into three key stages:

- **Feature Extraction Using BallMapper:** In this stage, financial attributes of companies are selected using the BallMapper algorithm, enabling a more nuanced evaluation of credit risk by capturing topological insights that go beyond conventional linear methods.
- **Graph Construction Using Extracted Features:** The extracted feature set is transformed into graphs by incorporating proximity relationships among companies, creating a non-Euclidean data structure suitable for graph analysis.
- **Credit Risk Prediction Using GNN:** Finally, the constructed graph is analyzed using a GNN model to predict the financial health of companies. This stage utilizes the GNN's message-passing framework to aggregate information from neighboring nodes and refine node representations for accurate classification.

The overall architecture of the proposed BM-GNN model is illustrated in Figure 1, which outlines the connections between these stages and highlights the flow of information. Algorithm 1 further details the steps involved in implementing the model, providing a comprehensive pseudocode representation of the feature extraction, graph construction, and prediction processes. Each of these stages will be elaborated in the following sections, offering a deeper understanding of the methods and rationale underlying the BM-GNN framework.

Algorithm 1 Pseudocode of BM-GNN Model

Input 1: Tabular companies' dataset of financial attributes X (sized $m \times n, i = 1, 2, \ldots, n$) and labels y_i (sized $m \times 1, i = 1, 2, \ldots, m$)
Input 2: Selection of an integer K
Input 3: Selection of *epsilon* as radius of the ball
Output: Predicted labels y'(sized $m \times 1, i = 1, 2, \ldots, m$)
for Co_i in all m companies: **do**
  Calculate Z_i-score
End for
Let tabular companies' dataset of financial attributes x_i as *points*, and Z_i-score as *values*
l = BallMapper (*points*, *value*, *epsilon*)
for all x_i in financial attributes: **do**
  ColorIgraphPlot(l)
End for
Stage I Completed Tabular companies' dataset of BallMapper attributes H (sized $m \times h, i = 1, 2, \ldots, n$)
for all x_i in H: **do**
  Compute Kendall's correlation: $k_{ij}(j = 1, 2, \ldots, h)$;
  Sort all the computed Kendall's values and select the k least instances as the neighborhood set of the instance $x_i : N_K(i)$
  for x_j in $N_K(i)$: **do**
    $a_{ij} = a_{ji} = 1$: Drow edge between x_i and x_j.
  end for
end for
Stage II Completed Transformed graph and its adjacency matrix obtained: $G_K, A_K = \left[a_{\mathrm{ij}}\right]$
for $G_K, A_K = \left[a_{\mathrm{ij}}\right]$ and node attributes X: **do**
  Feed into GNN and get renewed node representations $h^{(\ell+1)} = \delta\left(\widetilde{D}^{-\frac{1}{2}}\tilde{A}\widetilde{D}^{-\frac{1}{2}}h^{(\ell)}W^{(\ell)}\right)$[1];
  Update $W^{(\ell)}$ to optimize the losses between h_i^{ℓ} and y_i.
Stage III Completed Predicted labels is created $y' = [h_i^{\ell}]$

Stage I: Feature Selection Using BallMapper

As seen in Algorithm 1, the dataset comprises m companies, where X represents n features, and y denotes bankruptcy status. The primary objective revolves around binary classification, a task well suited for applying machine learning techniques on structured tabular data (Algorithm 1- Input 1). Initially, the focus lies in computing the probability of bankruptcy for each company. To achieve this, the Z-score model, a renowned methodology for credit scoring, is adopted, serving as the cornerstone of the study. (1) illustrates the discriminant function within the Altman's Z-score

[1] Where $h^{(\ell)}$ illustrates the node attribute matrix in layer ℓ, $W^{(\ell)}$ denotes the weight matrix for layer $\ell\tilde{D}$ is the degree matrix of $\tilde{A}$, $\tilde{A} = A + I$ is the adjacency matrix of the graph augmented with self-connections, and δ is the activation function.

model for industrial companies [44].

$$Z = 1.2\alpha_1 + 1.4\alpha_2 + 3.3\alpha_3 + 0.6\alpha_4 + 0.999\alpha_5 \tag{1}$$

where α_i is specified as follows for $i = 1, 2, 3, 4, 5$:

$$\alpha_1 = \frac{Working\ Capital}{Total\ Assets} = \frac{Current\ Assets - Current\ Liability}{Total\ Assets} \tag{2}$$

$$\alpha_2 = \frac{Retained\ Earnings}{Total\ Assets} \tag{3}$$

$$\alpha_3 = \frac{Earning\ Before\ Interest\ \&\ Taxes\ (EBIT)}{Total\ Assets} \tag{4}$$

$$\alpha_4 = \frac{Market\ Capitalisation}{Total\ Liabilities} \tag{5}$$

$$\alpha_5 = \frac{Sales}{Total\ Assets} \tag{6}$$

As Fig. 2 shows, if the Z-Score is above 3, it implies that the company is in a secure zone with minimal chances of facing financial difficulties. On the contrary, when the Z-score falls below 1.8, it triggers concerns, indicating financial distress and a significant likelihood of imminent bankruptcy. The range from 1.8 to 3 indicates a gray zone, suggesting a moderate risk of bankruptcy.

FIGURE 2. Altman's Z-scores zones.

Considering that this model computes company scores linearly, there is a possibility that it may not accurately predict the companies' situations. For instance, consider a company with low sales volume yet achieving maximum possible profits. Due to its low sales, this company would likely obtain a low score in the Altman model for credit scoring. To address this challenge, leveraging the BallMapper algorithm can be instrumental [45]. This algorithm rooted in topological data analysis, provides a powerful tool for representing high-dimensional datasets as point clouds. By employing a random selection of landmark points and adjusting representations through a configurable radius (epsilon), it ensures uniform and adaptable visualizations. BM graphs preserve intrinsic data relationships and structures, enabling the exploration of connectivity, clustering patterns, and topological features such as homology, even in dense datasets. This methodology facilitates the detection of complex patterns and correlations, offering nuanced insights into data that traditional techniques may overlook, particularly in analyzing corporate insolvency and financial behaviors.

As Algorithm 2 shows, the process of constructing a BM graph begins by selecting a high-dimensional dataset, denoted as X, and a radius *epsilon* to define the balls. Initially, all points in X are considered uncovered. The algorithm iterates through each uncovered point $p \in X$, marking the points within the ball of radius *epsilon* centered at p as covered.

Algorithm 2 Pseudocode of BallMapper Algorithm

Input 1: Let X be the high-dimensional dataset with point cloud X
Input 2: Selection of *epsilon* > 0 as radius of the ball
Output: BM graph
$C = \emptyset$;
$E = \emptyset$;
Let all point clouds of X as *uncovered*;
while There exist *uncovered* $p \in X$: **do**
  $C = C \cup p$;
  Mark all point $x \in \mathcal{B}(p, epsilon)$ as *covered*;
end
Define abstract vertices as V, which are extracted one per each element of C;
for $p_1, p_2 \in C | x \in X \cap \mathcal{B}(p_1, epsilon) \cap \mathcal{B}(p_2, epsilon)$: **do**
  $E = E \cup \{p_1, p_2\}$;
end

These covered points are then grouped into a cluster C. Once all points are assigned to clusters, abstract vertices V are defined, representing each cluster. The edges E of the graph are established by connecting two vertices if their corresponding clusters overlap, i.e., if there are points shared between the balls centered at p_1 and p_2 with radius *epsilon*. This results in a BM graph that captures the topological relationships between the clusters, revealing the underlying structure of the data. Figure 3 shows a sample BM graph generated by applying the BallMapper algorithm to a dataset of 250 data points with 40 features. This graph visually illustrates how the data points are organized according to their intrinsic characteristics, making it easier to identify key patterns and relationships within the dataset. In the following section, we will explore the main features of the BM graph and discuss how they enhance our understanding of the data.

Firstly, coloration is used to represent the average value of the default function within each ball, offering flexibility to users in selecting alternative functions such as counts, standard deviations, minimum, or maximum values. Sustainability scores are positioned on a scale on the right side of the plot, with lower scores situated towards the bottom and balls 9 and 22 averaging over 85.

Moreover, the size of balls reflects the number of entities within each ball, indicating data concentration. Larger balls signify denser concentrations of data points, as seen with ball 26 containing the most data while balls 25 and 2 have the lowest concentration. The number of balls provides insight into the intricacy of the analysis, with a greater number of balls required to encompass data points when decreasing the parameters of the ball radius. In constructing the graph depicted in Fig. 3, experiments were conducted with varying ball radii, ultimately determining a radius of $\varepsilon = 14.2$ as the optimal value. Connectivity among balls reveals overlap in data distribution, with numerous edges indicating similarity between data points. Smaller arms reaching individual balls suggest deviations from the central cluster, potentially indicating outliers as seen with balls 2 and 25. Lastly, correlation illustrates the relationship among axis variables, with correlated data points aligning along a narrow band. The

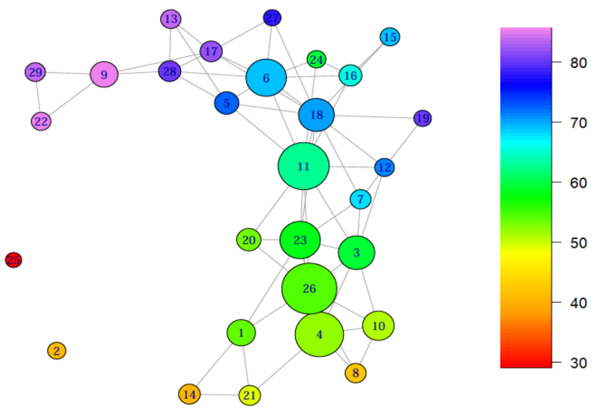


FIGURE 3. BallMapper graph sample (Note: This graph is related to a sustainability dataset of 250 companies in 2019 with 40 features. Coloration is the sustainability score. All axes are normalized to [0, 100]. $\varepsilon = 14.2$).

shape of the BM graph reflects this correlation, appearing more net-like when variables are less correlated, providing a conceptual representation of correlation within the data.

In our proposed model, we used the BM algorithm for feature selection. The reason for using BM is its ability to uncover the underlying topological structure of the data, which traditional methods might miss. By applying BM to the dataset, we can capture the geometric properties and the distances between data points (represented as balls in the graph). This method allows us to identify features that exhibit similar behavior to established credit risk measures, such as the Altman Z-score, and better reflect the complex, non-linear relationships within the data. To implement BM for feature selection, we utilized the BallMapper library in R, and as outlined in Algorithm 1, the function "ColorIgraphPlot" was employed to visualize the corresponding BM graph for the input data.

After applying BM to the labeled dataset using the Z-score model, a feature set, referred to as the **BM Feature Set**, is obtained. This feature set is then used as input for the next stage of analysis in our model.

Stage II: Graph Construction

The feature set derived from BallMapper encompasses m companies, with X representing attributes and y representing labels. The graphs constructed in this stage, characterized as a unique yet frequently encountered type of non-Euclidean structured data, comprise several elements, containing nodes (V), node features (X), edges (E), and bankruptcy status (y). Subsequently, the method for accessing these graphs is elaborated upon, detailing the procedure for obtaining access to these graph structures.

In the initial step, it is imperative to compute the correlation between each pair of companies based on the features selected by the BallMapper algorithm. Correlations are frequently utilized to capture interdependencies during typical market conditions. As an example, "Kendall's rank correlation coefficient" (Kendall's τ) serves as a "non-parametric" indicator of data associations, with values ranging from -1 to 1. Pozzi et al. [53] advocate for the use of Kendall's tau over Pearson correlation, contending that it provides richer information. In research conducted in [54], the comparison between Pearson correlations and two rank correlation techniques highlights the enhanced stability of rank methods, making them better suited for examining relationships in financial markets.

Suppose two observations (X_a, Y_a) and (X_b, Y_b) are two financial ratios of two companies. The coefficient τ can be define as:

$$\tau(a, b) = \Pr(\text{concordant}) - \Pr(\text{discordant}) \tag{7}$$

where probability of concordant moves of X and Y means $(X_a - X_b)(Y_a - Y_b) > 0$ and probability of discordant moves of X and Y means $(X_a - X_b)(Y_a - Y_b) < 0$.

Utilizing the aforementioned approach, we computed the correlation matrix between each pair of companies to lay the groundwork for graph plotting.

In the second step of this stage, we will derive an adjacency matrix for the graph using the proximity matrix obtained from the Kendall's correlation matrix. This adjacency matrix will represent the relationships between companies in the graph. This study introduces the application of the KNN algorithm to extract relational information from financial features, facilitating the construction of a network feature set essential for subsequent analysis. A fundamental representation of a graph, denoted as $G = (V, E)$, encompasses node features (X), bankruptcy status (y), and the adjacency matrix (A), which collectively define the graph constructures under investigation. The adjacency matrix A, typically symmetric and of size $m \times m$, signifies connections between nodes, with elements a_{ij} and a_{ji} represent the existence or absence of edges between nodes x_i and x_j. While in undirected unweighted graphs, a_{ij} and a_{ji} typically assume binary values (0 or 1), they may carry specific weights in weighted graphs. In the context of directed graphs, the matrix takes on an asymmetric form, introducing further complexities for analysis.

Node attributes and labels are succinctly represented by matrices X (of size $m \times n$) and y (of size $n \times 1$), respectively, where n signifies the dimensions of the attribute vectors. The establishment of graph structures necessitates the development of the adjacency matrix based on predefined criteria or rules. In BM-GNN model, companies are regarded as nodes, and K nearest neighbors is utilized to uncover connections between companies in an unsupervised manner. Through evaluating the Kendall's correlation of feature vectors among m companies, a ranked list of the closest neighbors for each node is determined. Determining K values in KNN governs the connections between nodes and their K closest neighbors. The resulting similarity measurements guide the construction of the adjacency matrix $A = [a_{\mathrm{ij}}]$, where $N_{G_K}(i)$ denotes the neighbor set of node i based on the chosen K value.

$$a_{\mathrm{ij}} = \begin{cases} 1, & x_i \in N_{G_K}(i) \wedge x_j \in N_{G_K}(j) \\ 0, & x_i \notin N_{G_K}(i) \vee x_j \notin N_{G_K}(j) \end{cases} \tag{8}$$

The resultant graphs, derived from the BallMapper algorithm and Kendall's correlation matrix, are characterized as undirected and unweighted, showcasing diverse edge configurations as delineated by distinct adjacency matrices. These graph representations furnish a robust data framework for subsequent predictive modeling stages. Noteworthy, the parameter K in our proposed model determines the number of neighbors of each node that are connected to it, thereby influencing the sparsity levels within the structed graphs. Consequently, K emerges as a pivotal yperparameter significantly influencing predictive performances.

In our proposed model, after calculating Kendall's correlation for all companies and constructing 8 adjacency matrices using the KNN model with K values of 5, 10, 15, 20, 25, 30, 35 and 40, we obtained 8 distinct graphs. By extracting features from these constructed graphs, we created 8 new feature sets, referred to as the **BM Network feature Set**. These feature sets are then used as input for the next stage of the analysis in our model.

Stage III: GNN Model

In this stage, a GNN is used to predict company bankruptcy. After the graph transformation and the construction of the adjacency matrix in stage II, the GNN processes the node features along with the graph structure to generate updated node representations. Each company is represented as a node, and its features are used as the node attributes. The adjacency matrix, derived from Kendall's correlation and the KNN-based graph, captures the relationships between the companies. The GNN performs a message-passing operation, where information is aggregated from neighboring nodes (i.e., other companies) based on their relationships in the graph.

The GNN model updates the node representations iteratively. In each layer ℓ, the node representations are updated using the following formula:

$$h^{(\ell+1)} = \delta\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}h^{(\ell)}W^{(\ell)}\right) \tag{9}$$

The updated node representations $h^{(\ell+1)}$ are then used to predict the bankruptcy status of each company. The bankruptcy prediction is made by comparing the final node representations after the last GNN layer with the true bankruptcy labels. The model optimizes the weight matrices $W^{(\ell)}$ during training to minimize the difference between the predicted labels and the true labels. This allows the model to predict whether a company is at risk of bankruptcy or not based on its features and the relationships within the graph.

IV. EXPERIMENTS

This section presents a thorough examination of the experiments carried out to validate the effectiveness of our proposed BM-GNN model. Our focus spans three key areas: data acquisition and preprocessing, feature selection and extraction, and results analysis alongside metrics for evaluating performance. The primary objective of this paper is to offer general insights into performance, facilitating clearer comparisons among various experimental methodologies, including the BM-GNN model, specifically in the domain of credit risk evaluation.

Up to this point, we have provided an explanation of the proposed model. It is important to note that to more accurately evaluate the performance of the proposed model, stages I and II were repeated for feature selection using the correlation relationships among the examined features. As a result, in stage I, a new feature set called the **Correlated Feature Set** was created and passed to stage II. Following the execution of stage II, a new feature set named the **Corr Network Feature Set** was generated.

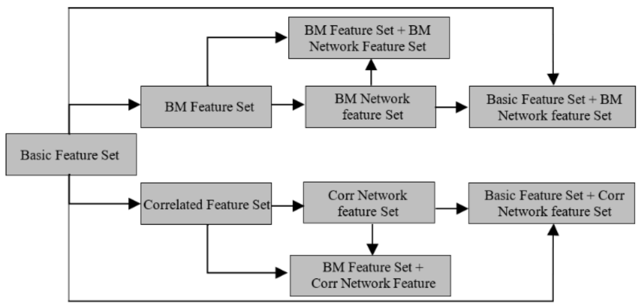


FIGURE 4. Extracted features and feature sets.

Finally, by evaluating different combinations of the created feature sets along with the **Basic Feature Set**, we input the generated datasets into the GNN model as well as four other machine learning models (Fig. 4). The results obtained from the evaluation of the implemented models are presented in the following section.

A. DATASET

The dataset used in this study consists of two information sources. The first source is the financial data platform of Iran markets, provided by Mofid Brokerage's Bourse View. The second source is the collection of financial statements of companies active in the energy sector, which were previously extracted from Mofid's Bourse View. As a result of aggregating this information, the financial indicators related to 2000 companies active in the energy sector, including oil and gas, petrochemicals, and electricity, are included. The companies in the raw dataset were classified according to their responsibilities and functions within the energy supply chain. The Primary Energy Producers (PEPs) are located at the highest level and are responsible for the exploration, extraction, and processing of natural energy resources. Tier 1 suppliers offer vital services and materials directly to PEPs, encompassing drilling equipment, pipelines, and sophisticated technical services. Tier 2 suppliers have expertise in providing certain technologies and services, including safety equipment, logistics, and specialist machinery, to assist Tier 1 providers. Tier 3 suppliers manufacture supplementary materials and components utilized by Tier 2 suppliers, including tools and fundamental construction supplies. Logistics and transportation firms oversee the transit of energy resources and products throughout the supply chain. Distributors procure processed energy commodities from manufacturers and distribute them to utility companies or directly to major customers, while retailers, such as local gas stations and energy service firms, disseminate these commodities to end consumers. In addition, technology suppliers supply software and hardware for energy management and exploration, while service providers offer maintenance, repair, and operating services. Each organization within the energy sector supply chain plays a crucial role in maximizing the efficiency of energy extraction, production, distribution, and maintenance activities.

The dataset is mainly centered around small and medium-sized enterprises (SMEs). It consists of 2000 companies carefully chosen from five distinct groups, each clearly labeled. There are 259 firms classified as Tier 1 suppliers, 491 companies classified as Tier 2 suppliers, 703 companies classified as Tier 3 suppliers, 147 companies classified as Technology providers, and 400 companies classified as Logistics and transportation companies. The dataset encompasses a wide array of enterprises linked to the energy industry, providing important insights into the financial behaviors, performance, and trends of the entities involved in the SCF. The information includes key financial indicators, such as revenue, expenses, profits, and other relevant financial metrics, enabling a detailed analysis of the economic dynamics within the energy sector supply chain. Before analysis, we performed data normalization to ensure that all features were on a comparable scale. This involved transforming the data so each feature had a mean of zero and a standard deviation of one. Normalization helps mitigate the impact of varying scales among features, enabling more effective analysis and modeling.

In this study, a specific indicator within the dataset determines the financial status of companies, distinguishing between bankruptcy and non-bankruptcy. The data indicates that approximately 70% of the companies are non-bankrupt, while 30% are classified as bankrupt. Given the diverse ML models employed in this paper, ensuring a balanced distribution of data between training and testing sets is imperative. To address this requirement, particularly in cases where data imbalance needed rectification, we adopted the methodology proposed in the Khemakhem study [46]. The ratio of bankrupt to non-bankrupt instances was systematically adjusted to achieve balance, adhering to varying proportions as specified in the cited reference.

B. FEATURE EXTRACTION

In this section, we will explore feature extraction methods aimed at reducing data dimensionality and improving model performance. Two approaches, the BallMapper algorithm and correlation-based extraction, are utilized to capture key patterns and relationships in the dataset.

1) USING BALLMAPPER ALGORITHM

In this stage, considering five financial indicators from the available set, encompassing "working capital to total assets", "accumulated retained earnings to total assets", "earning before interest and taxes (EBIT) to total assets", "market capitalization to total liabilities", and "sales to total assets", the Altman scoring model is implemented for each examined company [44]. Based on the implementation of this model, approximately 45% of companies were placed in the safe zone, around 25% of them in the gray zone, and 30% of them in the red zone. Considering that this model computes company scores linearly, there is a possibility that it may not accurately predict the companies' situations. To address this issue, we propose using the BallMapper method [45].

In order to draw a BallMapper graph, the optimal radius of the balls must first be obtained experimentally. In this aim, 30 BallMapper graphs were drawn with radii in the range [0, 3]. Then, using the "pointToBallList" function from the BallMapper package in R, we determined the number of balls drawn in each BallMapper graph and plotted the elbow plot with the aim of finding the optimal graph. Fig. 5 shows the number of balls created in the graph relative to various radii. As can be seen, radii in the range [1.5, 2] will likely create the best ball mper graph. Therefore, in order to obtain the optimal solution, they should be drawn and the optimal radius should be selected experimentally.

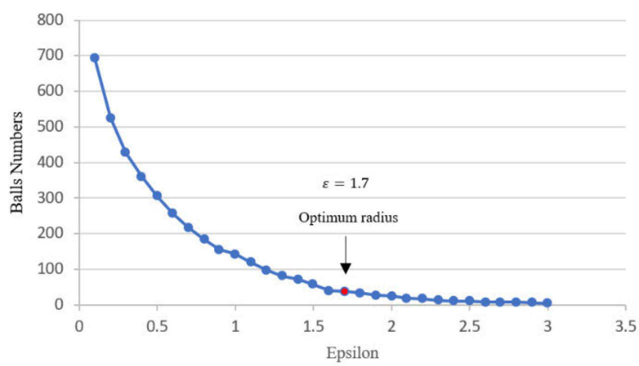


FIGURE 5. Ball radius and ball numbers.

For visual representation of radii and balls, Fig. 6 displays a total of 14 plots, enhancing readability and emphasizing differences within this range. It's important to highlight that the formation of the graph is contingent upon the data, thus discrepancies between graphs arise from variations in data distributions across the 93 axes, joint distributions, and the inclusion of differing numbers of observations.

Furthermore, Fig. 6 visually depicts the impact of a low ε parameter by demonstrating a value that results in an illegible diagram. Adjusting the parameter to $\varepsilon = 1.7$ in panel (i) yields a clearer plot, whereas setting $\varepsilon = 3$ in panel (n) does not yield significant insights into the space because of the restricted quantity of balls.

As discussed theoretically, there is no pre-established algorithm for determining the ideal number of balls, thus leaving this decision to the researcher's discretion. While constructing graphs similar to those depicted in Fig. 5 and 6 may offer some insights, the primary focus should be on the informativeness of the BallMapper graph as the key determinant.

Figure 7 illustrates a BallMapper graph with $\varepsilon = 1.7$, portraying a notable concentration of balls towards the center. The coloration is based on the average Z-score within each ball, representing a collection of companies exhibiting similar financial behavior from our dataset. A connection between two balls indicates the presence of at least one company shared between them. Given that this visualization aims to condense 93 dimensions into two-dimensional space, direct interpretation of the vertical or horizontal direction is not feasible. Nonetheless, the BM graph provides insights into the overall shape of the data; for more detailed examination of specific variables' behavior, coloration by those variables can be employed.

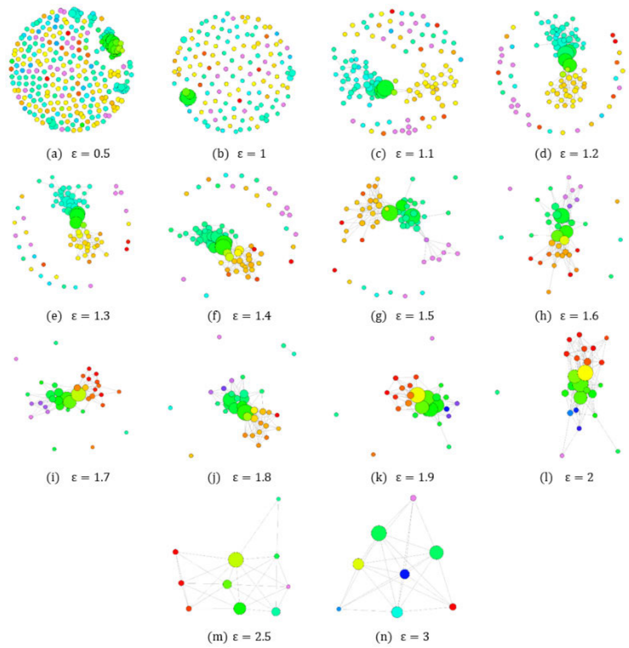


FIGURE 6. Different ball radius in BallMapper graphs.

As shown in Fig. 7, the balls positioned in the top-right corner of the chart represent a set of companies with the lowest Altman scores. However, as we approach the center, the average score increases, with the average score of companies within the balls at the center of the chart falling within the range of 2.5 to 3.5. Consequently, as we move away from the center of the chart, the average scores of companies within the balls increase, reaching an average higher than 4.5. It delineates how the distribution of companies corresponds to their Altman scores across the chart's spatial layout. By analyzing the spatial arrangement of the balls, one can discern trends in the distribution of Altman scores among the represented companies.

According to the characteristics of the BallMapper graph, in this study, we can utilize it to select financial ratios exhibiting similar behavior to the bankruptcy prediction model of Altman. To this end, we depict this graph for all financial ratios except for the five variables used in the Altman model.

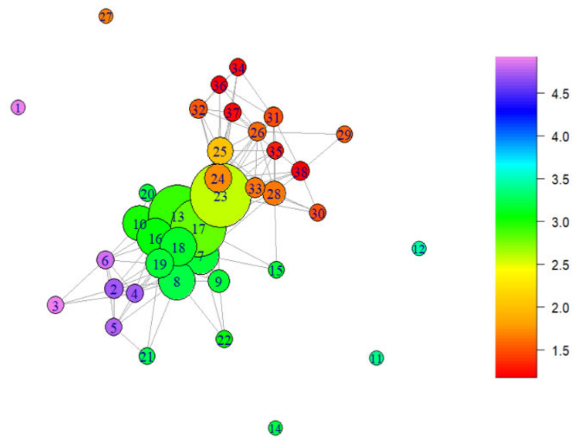


FIGURE 7. Altman Z-scores BM Graph(Note: This graph is related to a sustainability dataset of 2000 companies in 2022 with 93 features. Coloration is the Altman Z-scores. All axes are normalized to [0, 1]. $\varepsilon = 1.7$).

After examining the BallMapper diagrams for each of the 88 financial ratios, it was determined that 12 ratios emerged as the most influential factors the credit risk evaluation in SCF (Fig. 8).

Given the extensive number of input variables in the BallMapper algorithm utilized in this study, we have opted to selectively consider a subset of them for further analysis. We proceed by elucidating the methodology of financial ratio analysis for companies present within each cluster. This process illuminates the selection of 12 key indicators. It seems that there is an implicit assumption that the variables function independently. However, these variables may interact, affecting the results. One of the valuable features of the BallMapper model is that it can consider potential interactions between variables by displaying the color changes of the balls; if variables (a) and (b) interact with each other, the changes in color spectrum in the BM graph resulting from both variables will either be aligned with each other or be in opposite directions. Therefore, if variable (a is selected in feature selection using the BallMapper method, then variable (b) will also certainly be selected.

Investors and financial institutions may seek explanations for the significant variance in returns between firms covered by a specific ball and those adjacent to it. Additionally, the observation of low scores at the top of the plot in Fig. 7, despite similar axis variables, raises questions about why certain balls are not interconnected. Addressing these inquiries necessitates leveraging the ball comparison feature offered by Dłotko [47].

Here is illustrated the outcomes of comparisons between two distinct sets of balls derived from the BallMapper plot in Fig. 7.

From the given description, it's clear that the topmost ball, denoted as ball 1, is isolated in the upper portion of the space and not connected to any other balls. However, balls with similar average scores, precisely Balls 2, 3, 4, 5, and 6, have aggregated in the lower portion of Ball 1 in the space. As illustrated in Table. 1 these balls exhibit a higher "Inventory Turnover Ratio" compared to Ball 1. Remarkably, the Receivables "Turnover Ratio" of these companies is significantly lower in comparison to that of Ball 1. Upon normalization by the standard deviation, it becomes apparent that the primary distinction between Sphere 1 and the other Balls stems from the "Inventory Turnover Ratio", with the secondary variance arising from the Receivables Turnover Ratio. Expanding the comparison to encompass other isolated Balls specifically, incorporating Balls 11, 12, 14, and 27 into the central mass yields a similar trend. However, the gap in the "Inventory Turnover Ratio" narrows, leaving the "Receivables Turnover Ratio" as the probable determinant of differentiation. Notably, when employing a benchmark of a two-standard-deviation difference, none of the variables meet this criterion in the latter scenario.

In another example, we aim to investigate why Ball 24, despite being positioned adjacent to Ball 23, exhibits a lower volume and score compared to Ball 23. As depicted in Table. 1, Ball 24 demonstrates a lower "Inventory Turnover Ratio" relative to Ball 23. Furthermore, Ball 23's higher "Current Ratio" than Ball 24 indicates that it has more current assets relative to current liabilities, implying greater financial stability and better liquidity to meet short-term cash needs and debt obligations. Therefore, attention to these financial ratios can help identify potential reasons for the performance disparities between the two balls and facilitate a more detailed analysis of the factors influencing their respective performances.

Following the aforementioned procedure, we extracted the 12 most influential financial ratios by thoroughly examining all BM outputs associated with each financial ratio. Subsequently, as described in previous sections, by computing the Kendall correlation matrix on the BM feature set and implementing the KNN model with 8 K values (5, 10, 15, 20, 25, 30, 35, and 40), we obtained eight graphs. After constructing various graphs, we extracted eight network features for each graph to investigate the impact of graph properties on assessing credit risk for companies and predicting their short-term bankruptcy.

2) USING CORRELATION

This study performed a correlation examination to explore the relationships between the specified variables. Fig. 9 illustrates the relationships among the financial ratios, where darker shades of red signify stronger positive correlations, while deeper shades of blue indicate stronger negative correlations. Notably, out of the total 93 pairs, 53 pairs exhibit correlation absolute values exceeding 0.75, representing 56.98% of the total pairs. This observation underscores a pervasive strong correlation among most variables in the financing behavioral data. In this stage, considering a threshold of 0.75 in the calculations related to the correlation between financial ratios, 53 ratios were selected. Subsequently, by performing all the procedures described in the previous subsection, eight graphs were constructed with the same K values.

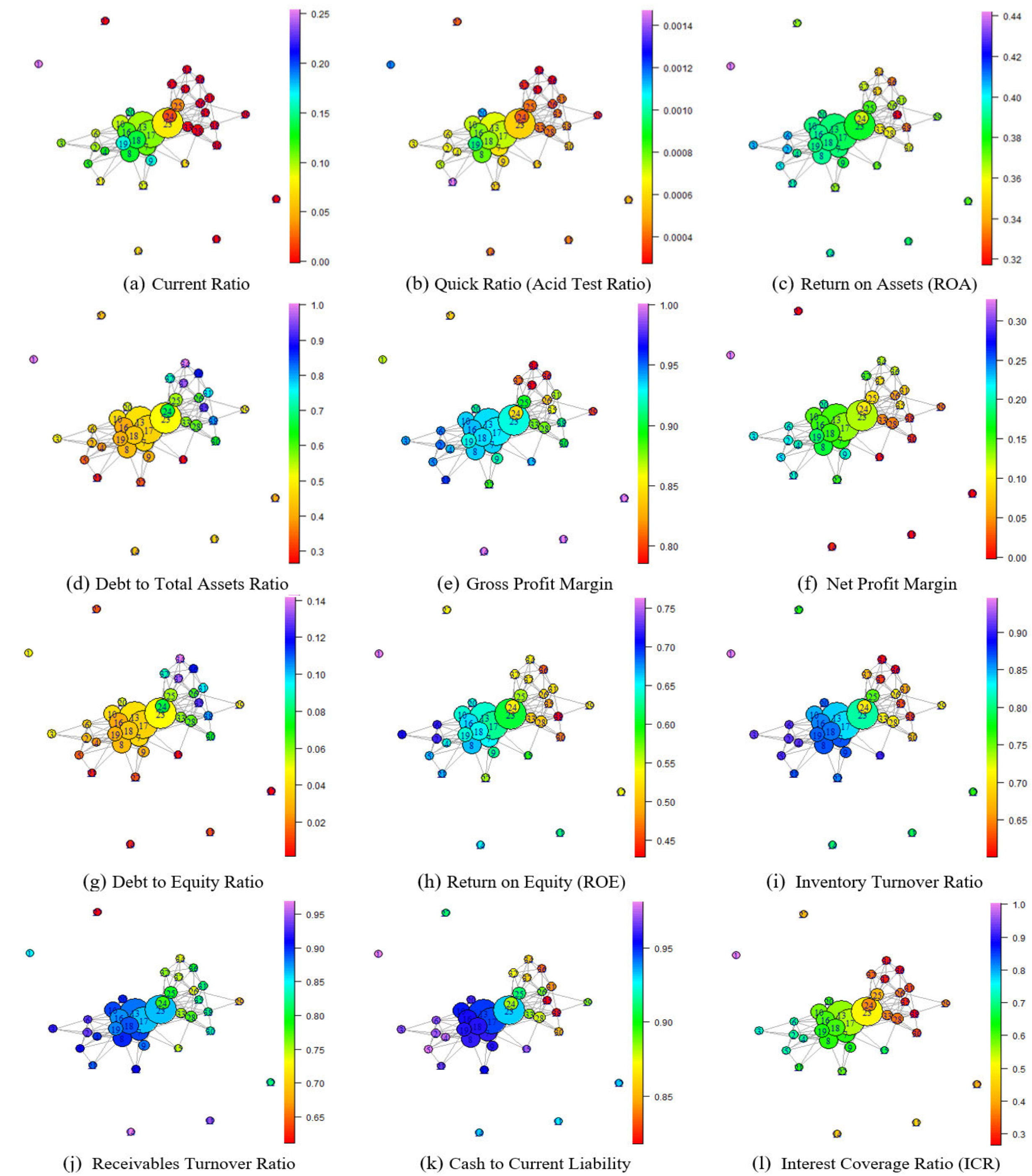
C. RESULTS

Utilizing the process described in the preceding section, this study obtained 18 distinct feature sets using various feature selection and extraction methods from the Basic Feature Set. Subsequently, by amalgamating the feature sets, we assess the efficiency of the GNN model against four ML models. To evaluate the performance of machine learning models and GNN, the dataset was randomly divided into two subsets: 80% for training data and 20% for testing data. To ensure class distribution preservation in both subsets, we used stratified sampling. This method guarantees that the class ratios in the training and testing subsets align with those in the original dataset, thus preventing biases in model evaluation.

The results of implementing ML and GNN models will be presented in nine scenarios as illustrated in Table. 2. The evaluation metrics utilized in the experiments are derived from well-established standard measures within the domain of credit risk prediction for SMEs in SCF. These metrics include average accuracy, precision rate, recall rate, F1-score rate, and the receiver operating characteristic (ROC) curve [48]. The limitations and challenges of evaluating GNNs using traditional metrics such as ROC curve are discussed in various research contexts [49], [50]. Therefore, in the following sections, the AUC value derived from the ROC curve is calculated only for evaluating the ML models. For the GNN model evaluation, this metric could not be computed.

– Scenario 1: In this scenario, the results obtained by employing four different machine learning models on the "Basic Feature Set", which includes 93 financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

Table. 3 shows the accuracy rates of LR, DT, MLP, and SVM models in assessing credit risk for companies, all achieving 90.50 percent accuracy except SVM, which lags by 0.75 percent. DT outperforms others with an F1-score of 56.82 percent, indicating balanced performance. LR, MLP, and SVM models exhibit the highest AUC on the ROC curve, suggesting their superior ability to differentiate bankrupt and non-bankrupt companies. DT emerges as the most effective model for predicting SME bankruptcy in the energy sector SC. However, generalization to other industries or company types should be approached cautiously.



(a) Current Ratio (b) Quick Ratio (Acid Test Ratio) (c) Return on Assets (ROA)

(d) Debt to Total Assets Ratio (e) Gross Profit Margin (f) Net Profit Margin

(g) Debt to Equity Ratio (h) Return on Equity (ROE) (i) Inventory Turnover Ratio

(j) Receivables Turnover Ratio (k) Cash to Current Liability (l) Interest Coverage Ratio (ICR)

FIGURE 8. Results of BallMapper feature extraction.

In Table. 4, GNN implementation on feature selection via the BallMapper algorithm outperforms correlation-based feature selection, notably achieving its highest accuracy at $K = 25$, reaching 92.92 percent. The BallMapper algorithm's feature selection, renowned for enhancing interpretability and predictive accuracy, significantly impacts GNN's predictive capabilities. Furthermore, the BM model attains a remarkable F1-score of 92.48 percent at $K = 25$. A key observation is the smaller numerical differences between evaluation metrics in GNN compared to ML models, indicating superior credit risk evaluation performance in SCF using GNN. The utilization of graph-based techniques such as BallMapper further enhances the efficacy of GNN in extracting meaningful insights from complex data structures.

TABLE 1. Ball comparisons for Altman Z-score.

Compare: / With:	Ball 1 / Balls 2, 3, 4, 5, 6		Ball 23 / Ball 24	
	Dist.	**Diff.**	**Dist.**	**Diff.**
Current Ratio	-0.394	0.810	-0.489	1.005
Quick Ratio (Acid Test Ratio)	0.692	0.265	2.396	0.916
Return on Assets (ROA)	0.061	0.365	0.016	0.099
Debt to Total Assets Ratio	1.164	0.806	0.450	0.311
Gross Profit Margin	-1.420	-0.334	0.450	0.029
Net Profit Margin	1.163	2.019	-0.015	-0.026
Debt to Equity Ratio	0.026	0.002	0.063	0.004
Return on Equity (ROE)	0.172	0.013	0.047	0.003
Inventory Turnover Ratio	-12.220	-5.959	-1.251	0.610
Receivables Turnover Ratio	10.274	3.482	0.211	0.071
Cash to Current Liability	0.209	0.011	0.009	0.000
Interest Coverage Ratio (ICR)	0.945	0.536	0.667	0.378

TABLE 2. Ball comparisons for Altman Z-score.

	Experiments	**Relevant Scenario**	**Description of the Scenario**	**Number of Features**
1	Basic Feature Set	1	Implementing models on a feature set, which contains all financial ratios of companies.	93
2	BM Feature Set	2	Implementing models on a feature set, which contains selected ratios using the BallMapper algorithm.	12
3	Correlated Feature Set	3	Implementing models on a feature set, which contains selected ratios using inter-company correlation relationships.	53
4	BM Network feature Set	4	Implementing models on a feature set, which contains network features constructed based on the BallMapper algorithm.	8
5	Corr Network feature Set	5	Implementing models on a feature set, which contains network features constructed based on inter-company correlation relationships.	8
6	Basic Feature Set + BM Network feature Set	6	Implementing models on a feature set, which contains the total of all financial ratios and network features constructed based on the BallMapper algorithm.	101
7	Basic Feature Set + Corr Network feature Set	7	Implementing models on a dataset comprising the total of all financial ratios and network features constructed based on inter-company correlation relationships.	101
8	BM Feature Set + BM Network Feature Set	8	Implementing models on a feature set, which contains selected ratios using the BallMapper algorithm and constructed network features.	20
9	Correlated Feature Set + Corr Network feature Set	9	Implementing models on a feature set, which contains selected ratios using inter-company correlation relationships and constructed network features.	61

– Scenario 2: In this scenario, the results obtained by employing four different machine learning models on the ‘‘BM Feature Set’’, which includes 12 financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

As shown in Table. 5, The evaluation of various machine learning models for credit risk assessment in supply chain finance reveals intriguing insights. Among the models studied, LR, MLP, and SVM demonstrate comparable accuracy at 89.75%. However, the DT model lags slightly behind with a 6.50% difference. Interestingly, although the DT model's accuracy is lower, its F1- score surpasses that of the other models at 27.96%, indicating a superior balance between precision and recall. Further analysis delves into the area

FIGURE 9. Correlations between the variables of the financial indicators of the companies.

TABLE 3. Evaluation of machine learning models in Scenario 1 (percent).

Model	F1-score	Recall	Precision	Accuracy	AUC
LR	17.39	9.76	80.00	**90.50**	**0.87**
DT	56.82	60.98	53.19	**90.50**	0.76
MLP	34.48	24.39	58.82	**90.50**	**0.87**
SVM	9.17	5.00	55.00	89.75	**0.87**

under the curve, where LR and SVM models outshine the others with a score of 0.73. This implies stronger discriminatory power in distinguishing between positive and negative instances, enhancing their overall effectiveness in predicting credit risk. Moreover, the study reveals nuanced performance nuances concerning false positives. While SVM performs best in controlling false positives until the rate reaches 0.4, LR outperforms beyond this threshold. This suggests the importance of selecting models based on specific risk tolerance levels.

As shown in Table. 5, The evaluation of various machine learning models for credit risk assessment in supply chain finance reveals intriguing insights. Among the models studied, LR, MLP, and SVM demonstrate comparable accuracy at 89.75%. However, the DT model lags slightly behind with a 6.50% difference. Interestingly, although the DT model's accuracy is lower, its F1- score surpasses that of the other models at 27.96%, indicating a superior balance between precision and recall. Further analysis delves into the area under the curve, where LR and SVM models outshine the others with a score of 0.73. This implies stronger discriminatory power in distinguishing between positive and negative instances, enhancing their overall effectiveness in predicting credit risk. Moreover, the study reveals nuanced performance nuances concerning false positives. While SVM performs best in controlling false positives until the rate reaches 0.4, LR outperforms beyond this threshold. This suggests the importance of selecting models based on specific risk tolerance levels.

As indicated in Table. 6, the implemented GNN model achieved the highest accuracy on the feature set resulting from Implementation of BallMapper algorithm, specifically with K = 20 reaches 92.91 percent. This signifies the robust performance of the GNN model in accurately evaluating credit risk. Furthermore, considering the F1-score, it becomes evident that this GNN model outperforms others by a margin of more than one percent, highlighting its superiority in achieving the highest value of this metric among the models. These findings underscore the efficacy of the GNN model, particularly concerning the feature selection process conducted by the BallMapper algorithm, for precise and reliable prediction of credit risk.

– Scenario 3: In this scenario, the results obtained by employing four different machine learning models on the "Correlated Feature Set", which includes 53 financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

As shown in Table. 7, implementing four ML models on the selected indices using correlation calculations reveals that both LR and DT models exhibit the highest accuracy on this dataset. Furthermore, the DT model stands out significantly in the F1-score metric compared to other models, occupying the highest tier with a substantial margin, achieving a value of 53.01%. These findings underscore the superior performance of LR and DT models in accurately assessing the feature set, with the DT model provides a superior balance between precision and recall in evaluating credit risk.

The MLP model outperforms the other three models, showcasing its superior performance. This suggests that MLP exhibits stronger discriminatory power in distinguishing between positive and negative instances when evaluating credit risk. The elevated AUC value for the MLP model implies better overall effectiveness and performance in predicting credit risk compared to the other models.

The results obtained from this section demonstrate that although the accuracy of the LR and DT models is higher compared to the other two algorithms, SVM achieves the highest AUC value. Therefore, considering all evaluation metrics, the MLP model, despite having lower accuracy of 0.25 compared to the LR and DT models, exhibits better performance. However, selecting the best model in these cases should be done in accordance with the business management approaches.

As indicated in Table. 8, the implemented GNN model achieved the highest accuracy on K = 35 with 90.27 percent. This signifies the robust performance of the GNN model in accurately evaluating credit risk. An important observation in comparing the results obtained from implementing the GNN model on the feature set derived from feature selection using correlation with the GNN model implemented in the previous scenario is that the accuracy in this section is 2.64% higher

TABLE 4. Evaluation of GNN model in Scenario 1 (percent).

K-Values	BM Feature Set				Correlated Feature Set			
	F1 score	Recall	Precision	Accuracy	F1 score	Recall	Precision	Accuracy
K = 5	88.41	89.39	87.46	89.57	86.6	88.51	84.78	88.78
K = 10	90.67	91.43	89.92	91.47	87.78	88.82	86.76	89.03
K = 15	90.6	91.23	89.97	91.41	88.85	89.8	87.92	**89.87**
K = 20	91.06	91.51	90.61	91.68	87.25	88.45	86.08	88.77
K = 25	92.48	92.85	92.11	**92.92**	87.67	88.69	86.67	88.85
K = 30	90.87	91.27	90.47	91.5	88.89	88.88	88.91	89.26
K = 35	90.91	90.6	91.23	90.63	89.08	89.54	88.63	89.81
K = 40	90.19	90.34	90.04	90.36	88.33	89.34	87.34	89.82

TABLE 5. Evaluation of machine learning models in Scenario 2 (percent).

Model	F1-score	Recall	Precisio n	Accuracy	AU C
LR	16.48	17.00	16.00	**89.75**	**0.73**
DT	27.96	31.71	25.00	83.25	0.69
MLP	6.67	10.00	5.00	**89.75**	0.70
SVM	17.14	20.00	15.00	**89.75**	**0.73**

TABLE 6. Evaluation GNN model in Scenario 2 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K = 5$	89.99	90.21	89.78	90.56
$K = 10$	90.66	90.68	90.64	90.93
$K = 15$	90.45	90.73	90.18	90.92
$K = 20$	92.58	92.86	92.3	**92.91**
$K = 25$	91.31	91.77	90.85	91.8
$K = 30$	91.41	91.68	91.14	92.1
$K = 35$	90.43	90.93	89.94	91.16
$K = 40$	90.76	91.17	90.36	91.42

TABLE 7. Evaluation of machine learning models in Scenario 3 (percent).

Model	F1-score	Recall	Precisio n	Accuracy	AU C
LR	17.02	17.02	66.67	**90.25**	0.81
DT	53.01	53.01	52.38	**90.25**	0.74
MLP	31.03	39.13	52.94	90.00	**0.85**
SVM	10.00	8.00	10.00	89.75	0.81

than in the previous scenario. This indicates that the BM model has positively influenced the performance improvement of the GNN model.

TABLE 8. Evaluation GNN model in Scenario 3 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K = 5$	87.54	88.91	86.21	88.95
$K = 10$	87.58	89.16	86.06	89.66
$K = 15$	87.54	88.75	86.36	89.12
$K = 20$	85.87	88.58	83.32	88.72
$K = 25$	85.17	88.17	82.37	88.51
$K = 30$	89.16	90.07	88.27	90.21
$K = 35$	88.78	89.84	87.74	**90.27**
$K = 40$	87.04	88.88	85.28	89.28

– Scenario 4: In this scenario, the results obtained by employing four different machine learning models on the "BM Network Feature Set", which includes eight features, are examined. Subsequently, the results generated by the GNN model are discussed.

The evaluation of machine learning models on the BM Network feature set reveals interesting insights (Table. 9). LR, MLP, and SVM achieved the highest accuracy of 91.50%. Notably, LR with K values of 5, 15, and 20, along with MLP and SVM with K values of 15 and 20, demonstrated the highest accuracy. However, despite the balanced nature of the data, all examinations on this feature set consistently yielded the highest F1-score value of 41.27%, indicating potential mismatch between the models and the features. Analyzing the ROC curves across different K values, MLP excelled with an AUC of 0.65 at K=5, while DT outperformed others with AUC values of 0.68 at K=10 and 0.79 at K=20,25, and 40. This suggests DT's superior performance across various thresholds.

Overall, despite its lower accuracy compared to other models, DT consistently performs better, as indicated by its AUC values and positioning within the ROC space.

Table. 10 reveals a notable consistency in performance metrics, including "F1-score", "Recall", "Precision", and

TABLE 9. Evaluation of machine learning models in Scenario 4 (percent).

	K-Values	F1-score	Recall	Precision	Accuracy	AUC		K-Values	F1-score	Recall	Precision	Accuracy	AUC
Logistic Regression	$K = 5$	32.00	19.51	88.89	**91.50**	0.63	Multi-layer Perceptron	$K = 5$	28.57	17.07	87.50	91.25	0.65
	$K = 10$	28.57	17.07	87.50	91.25	0.66		$K = 10$	28.57	17.07	87.50	91.25	0.67
	$K = 15$	29.17	17.07	100.00	**91.50**	0.73		$K = 15$	29.17	17.07	100.00	**91.50**	0.73
	$K = 20$	32.00	19.51	88.89	**91.50**	0.74		$K = 20$	32.00	19.51	88.89	**91.50**	0.75
	$K = 25$	32.73	21.95	64.29	90.75	0.68		$K = 25$	16.67	9.76	57.14	90.00	0.67
	$K = 30$	23.08	14.63	54.55	90.00	0.77		$K = 30$	23.08	14.63	54.55	90.00	0.75
	$K = 35$	27.45	17.07	70.00	90.75	0.69		$K = 35$	17.78	9.76	100.00	90.75	0.68
	$K = 40$	29.63	19.51	61.54	90.50	0.66		$K = 40$	21.74	12.20	100.00	91.00	0.60
Decision Trees	$K = 5$	26.42	17.07	58.33	90.25	0.63	Support Vector Machines	$K = 5$	17.78	9.76	100.00	90.75	0.52
	$K = 10$	38.60	26.83	68.75	**91.25**	0.68		$K = 10$	25.53	14.63	100.00	91.25	0.58
	$K = 15$	38.71	29.27	57.14	90.50	0.73		$K = 15$	29.17	17.07	100.00	**91.50**	0.68
	$K = 20$	29.63	19.51	61.54	90.50	0.79		$K = 20$	32.00	19.51	88.89	**91.50**	0.59
	$K = 25$	33.90	24.39	55.56	90.25	0.74		$K = 25$	32.73	21.95	64.29	90.75	0.62
	$K = 30$	40.63	31.71	56.52	90.50	0.80		$K = 30$	37.29	26.83	61.11	90.75	0.72
	$K = 35$	41.27	31.71	59.09	90.75	0.79		$K = 35$	27.45	17.07	70.00	90.75	0.57
	$K = 40$	24.00	14.63	66.67	90.50	0.72		$K = 40$	21.74	12.20	100.00	91.00	0.66

"Accuracy", which are observed to fall within a narrow range, spanning from 87.90 percent to 92.38 percent. Such stability reflects a marked difference from conventional machine learning model implementations, highlighting a distinctive trait of the GNN model. Notably, the assessment of credit risk attains its highest accuracy at K = 20, reaching 92.02 percent.

TABLE 10. Evaluation GNN model in Scenario 4 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K = 5$	90.15	89.42	90.89	89.86
$K = 10$	88.82	87.90	89.75	88.40
$K = 15$	91.12	90.50	91.74	90.90
$K = 20$	92.12	91.86	92.38	**92.02**
$K = 25$	91.28	90.75	91.82	91.00
$K = 30$	91.64	91.12	92.16	91.41
$K = 35$	90.45	90.87	90.04	91.36
$K = 40$	90.15	89.41	90.91	89.86

– **Scenario 5:** In this scenario, the results obtained by employing four different machine learning models on the "Correlated Network Feature Set", which includes eight features, are examined. Subsequently, the results generated by the GNN model are discussed. Table. 11 illustrates the evaluation of machine learning models, where LR achieves the highest accuracy at 89.75%. However, when considering F1-scores, SVM and DT models with a graph generated using K = 25 perform optimally, scoring 20.31% and 25.29% respectively. This indicates a trade-off between accuracy and F1-score, with LR excelling in accuracy while DT shines in F1-score. Analyzing ROC curves across different K values, LR consistently leads in performance, particularly at K = 5, 10, 15, 20, and 25. However, MLP outperforms at K = 30, and LR again takes the lead at K = 35 and 40.

Overall, the performance of ML models on this feature set is deemed subpar. Nevertheless, LR and MLP demonstrate superior performance compared to SVM and DT, especially noticeable at K = 30. Further examination is warranted to determine the best-performing model based on additional criteria.

Table. 12 demonstrates a remarkable consistency across various performance metrics, including "F1-score", "Recall", "Precision," and "Accuracy", all of which exhibit a narrow range from 76.74 percent to 89.76 percent. This consistency signifies a departure from typical machine learning model implementations, underscoring a distinctive characteristic of the GNN model. Particularly noteworthy is the highest accuracy achieved in assessing credit risk at K = 30, reaching 89.67 percent.

– **Scenario 6:** In this scenario, the results obtained by employing four different machine learning models on the "Basic Feature Set + BM Network Feature Set", which includes 101 network features and financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

In Table. 13, LR and DT models achieve the highest accuracy rate at 93.55%. A closer examination of F1-scores

TABLE 11. Evaluation of machine learning models in Scenario 5 (percent).

	K-Values	F1-score	Recall	Precision	Accuracy	AUC		K-Values	F1-score	Recall	Precision	Accuracy	AUC
Logistic Regression	$K = 5$	$K = 5$	7.21	5.23	11.60	88.75	Multi-layer Perceptron	$K = 5$	8.25	6.23	12.20	83.51	0.49
	$K = 10$	$K = 10$	12.88	20.00	9.50	89.73		$K = 10$	13.59	19.24	10.50	80.05	0.42
	$K = 15$	$K = 15$	18.38	20.00	17.00	88.51		$K = 15$	12.01	12.21	11.82	79.25	0.51
	$K = 20$	$K = 20$	12.30	8.78	20.54	89.56		$K = 20$	17.03	10.78	40.54	**87.65**	0.59
	$K = 25$	$K = 25$	21.35	19.56	23.50	87.82		$K = 25$	20.10	17.56	23.50	83.75	0.57
	$K = 30$	$K = 30$	11.76	19.23	8.47	88.65		$K = 30$	10.87	11.28	10.49	83.00	0.64
	$K = 35$	$K = 35$	12.20	14.68	10.43	88.23		$K = 35$	13.87	20.68	10.43	82.75	0.61
	$K = 40$	$K = 40$	14.02	15.07	13.11	**89.75**		$K = 40$	15.04	15.75	14.40	87.52	0.55
Decision Trees	$K = 5$	$K = 5$	10.53	9.76	11.43	83.00	Support Vector Machines	$K = 5$	21.57	20.00	23.40	87.24	0.54
	$K = 10$	$K = 10$	14.89	17.07	13.21	80.00		$K = 10$	14.93	18.45	12.54	86.52	0.64
	$K = 15$	$K = 15$	12.24	14.63	10.53	78.50		$K = 15$	14.61	15.87	13.54	88.26	0.49
	$K = 20$	$K = 20$	14.46	14.63	14.29	82.25		$K = 20$	20.98	20.24	21.78	87.52	0.56
	$K = 25$	$K = 25$	25.29	26.83	23.91	83.75		$K = 25$	20.31	18.49	22.53	88.72	0.48
	$K = 30$	$K = 30$	16.28	17.07	15.56	82.00		$K = 30$	18.28	21.54	15.87	88.22	0.41
	$K = 35$	$K = 35$	12.50	12.20	12.82	82.50		$K = 35$	13.59	14.68	12.65	88.53	0.49
	$K = 40$	$K = 40$	15.79	14.63	17.14	**84.00**		$K = 40$	15.61	15.75	15.48	**89.01**	0.46

TABLE 12. Evaluation GNN model in Scenario 5 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K = 5$	83.85	89.02	79.24	89.02
$K = 10$	82.91	88.28	78.15	88.40
$K = 15$	82.34	87.82	77.51	88.04
$K = 20$	81.81	87.60	76.74	87.60
$K = 25$	83.04	88.14	78.50	88.60
$K = 30$	84.61	89.28	80.41	**89.76**
$K = 35$	82.74	88.09	78.01	88.33
$K = 40$	84.60	89.41	80.29	89.61

reveals that the DT model, particularly with a graph generated using K = 40, attains the highest values for both accuracy and F1-score metrics, establishing itself as the superior model for credit risk assessment on this dataset. Analyzing ROC curves across different K values, MLP outperforms other models at K = 5, while LR and MLP surpass others at K = 10, and LR leads at K = 15. MLP showcases superior performance at K = 20, 30, and 40. These results highlight the robustness and effectiveness of LR and MLP models in capturing financial dataset patterns and nuances, underscoring their potential for accurate and reliable financial prediction tasks across various scenarios and parameter settings.

Table. 14 reveals a notable consistency in performance metrics, including "F1-score", "Recall", "Precision", and "Accuracy", which are observed to fall within a narrow range, spanning from 88.04 percent to 93.56 percent. Such stability reflects a marked difference from conventional machine learning model implementations, highlighting a distinctive trait of the GNN model. Notably, the assessment of credit risk attains its highest accuracy at K = 20, reaching 93.56 percent.

– Scenario 7: In this scenario, the results obtained by employing four different machine learning models on the "Basic Feature Set + Correlated Network Feature Set", which includes 101 network features and financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

Table. 15 presents the implementation of ML models, where SVM achieves the highest accuracy at 91.25%. However, upon detailed examination of F1-scores, the DT model, particularly with a graph generated using K = 5, demonstrates optimal performance with an F1-score of 60.46%. Further analysis delves into ROC curves across varying graph complexities (determined by K values), enhancing understanding of how these complexities impact ML model performance. At K = 5, both MLP and SVM models outperform others with an AUC of 0.89. LR leads at K = 10, while MLP performs better at K = 15, 20, 25, 30, 35, and 40, collectively demonstrating stronger performance. The LR model's accuracy remains relatively stable despite changes in K, showcasing consistent performance compared to other models. However, business conditions should be considered alongside model selection for credit risk assessment in the supply chain, emphasizing the importance of aligning ML model choices with specific business requirements.

TABLE 13. Evaluation of machine learning models in Scenario 6 (percent).

	K-Values	F1-score	Recall	Precision	Accuracy	AUC		K-Values	F1-score	Recall	Precision	Accuracy	AUC
Logistic Regression	$K = 5$	49.12	34.15	87.50	92.75	0.91	Multi-layer Perceptron	$K = 5$	52.31	41.46	70.83	92.25	0.92
	$K = 10$	51.72	36.59	88.24	93.00	0.88		$K = 10$	47.46	34.15	77.78	92.25	0.88
	$K = 15$	51.72	36.59	88.24	93.00	0.92		$K = 15$	47.46	34.15	77.78	92.25	0.91
	$K = 20$	54.24	39.02	88.89	93.25	0.90		$K = 20$	55.88	46.34	70.37	92.50	0.93
	$K = 25$	48.39	36.59	71.43	92.00	0.92		$K = 25$	54.55	43.90	72.00	92.50	0.93
	$K = 30$	55.38	43.90	75.00	92.75	0.90		$K = 30$	54.55	43.90	72.00	92.50	0.92
	$K = 35$	58.46	46.34	79.17	93.25	0.90		$K = 35$	60.00	51.22	72.41	93.00	0.90
	$K = 40$	61.54	48.78	83.33	**93.55**	0.91		$K = 40$	60.61	48.78	80.00	**93.50**	0.93
Decision Trees	$K = 5$	58.14	60.98	55.56	91.00	0.76	Support Vector Machines	$K = 5$	17.78	9.76	100.00	90.75	0.75
	$K = 10$	71.11	78.05	65.31	93.50	0.71		$K = 10$	25.53	14.63	100.00	91.25	0.71
	$K = 15$	62.92	68.29	58.33	91.75	0.78		$K = 15$	29.17	17.07	100.00	91.50	0.78
	$K = 20$	59.26	58.54	60.00	91.75	0.79		$K = 20$	32.00	19.51	88.89	91.50	0.72
	$K = 25$	55.17	58.54	52.17	90.25	0.76		$K = 25$	32.73	21.95	64.29	90.75	0.78
	$K = 30$	64.44	70.73	59.18	92.00	0.81		$K = 30$	26.42	17.07	58.33	90.25	0.79
	$K = 35$	69.77	73.17	66.67	93.50	0.79		$K = 35$	32.65	19.51	100.00	91.75	0.74
	$K = 40$	73.56	78.05	69.57	**93.55**	0.84		$K = 40$	36.00	21.95	100.00	**92.00**	0.80

TABLE 14. Evaluation GNN model in Scenario 6 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K = 5$	88.90	89.78	88.04	89.79
$K = 10$	89.93	90.42	89.45	90.74
$K = 15$	90.65	90.58	90.72	90.99
$K = 20$	93.17	93.29	93.06	**93.56**
$K = 25$	91.93	92.09	91.77	92.51
$K = 30$	91.14	91.29	90.99	91.71
$K = 35$	91.86	91.94	91.79	92.19
$K = 40$	90.68	90.98	90.38	91.37

Table. 16 reveals a notable consistency in performance metrics, including "F1-score", "Recall", "Precision", and "Accuracy", which are observed to fall within a narrow range, spanning from 85.11 percent to 91.51 percent. Such stability reflects a marked difference from conventional machine learning model implementations, highlighting a distinctive trait of the GNN model. Notably, the assessment of credit risk attains its highest accuracy at K = 30, reaching 91.51 percent.

– Scenario 8: In this scenario, the results obtained by employing four different machine learning models on the "BM Feature Set + BM Network Feature Set", which includes 20 network features and financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

In Table. 17, MLP achieves the highest accuracy among the implemented ML models on the feature set, particularly when combining BM features and network features obtained with K = 5. However, LR and DT models attain a similar accuracy of 91.75 under similar conditions. Notably, MLP also demonstrates superior F1-score performance compared to the other models at K = 5. Analyzing ROC curves across different K values, LR and MLP consistently outperform other models at K = 10, 20, and 25, with AUC values of 0.80, 0.81, and 0.83 respectively. MLP leads at K = 15 with an AUC of 0.81, while LR dominates at K = 5, 30, 35, and 40, showcasing AUC values of 0.76, 0.86, 0.85, and 0.84 respectively, indicating its superior positioning within the ROC space compared to alternative models.

Table. 18 reveals a notable consistency in performance metrics, including "F1-score", "Recall", "Precision", and "Accuracy", which are observed to fall within a narrow range, spanning from 87.86 percent to 92.46 percent. Such stability reflects a marked difference from conventional machine learning model implementations, highlighting a distinctive trait of the GNN model. Notably, the assessment of credit risk attains its highest accuracy at K = 30, reaching 92.46 percent.

– Scenario 9: In this scenario, the results obtained by employing four different machine learning models on the "Correlated Feature Set + Correlated Network Feature Set", which includes 61 network features and financial ratios, are examined. Subsequently, the results generated by the GNN model are discussed.

TABLE 15. Evaluation of machine learning models in Scenario 7 (percent).

	K-Values	F1-score	Recall	Precision	Accuracy	AUC		K-Values	F1-score	Recall	Precision	Accuracy	AUC
Logistic Regression	$K=5$	17.03	9.76	66.67	90.25	0.84	Multi-layer Perceptron	$K=5$	40.00	31.71	54.17	90.25	0.89
	$K=10$	17.03	9.76	66.67	90.25	0.85		$K=10$	23.52	14.63	60.00	90.25	0.87
	$K=15$	20.84	12.20	71.43	**90.50**	0.87		$K=15$	35.09	24.39	62.50	**90.45**	0.88
	$K=20$	21.28	12.20	83.33	**90.50**	0.89		$K=20$	26.41	17.07	58.33	90.25	0.89
	$K=25$	17.40	9.76	80.00	**90.50**	0.86		$K=25$	22.64	14.63	50.00	89.75	0.87
	$K=30$	17.03	9.76	66.67	90.25	0.90		$K=30$	29.09	19.51	57.14	90.25	0.90
	$K=35$	17.40	9.76	80.00	**90.50**	0.90		$K=35$	23.52	14.63	60.00	90.25	0.91
	$K=40$	17.40	9.76	80.00	**90.50**	0.88		$K=40$	33.33	24.39	52.63	90.00	0.89
Decision Trees	$K=5$	60.46	63.41	57.78	90.50	0.78	Support Vector Machines	$K=5$	17.78	9.76	100.00	90.75	0.89
	$K=10$	54.54	58.54	51.06	90.00	0.74		$K=10$	25.53	14.63	100.00	**91.25**	0.87
	$K=15$	57.98	63.41	53.41	90.48	0.74		$K=15$	40.00	31.71	54.17	90.25	0.87
	$K=20$	58.14	60.98	55.56	**91.00**	0.74		$K=20$	45.37	40.54	51.51	90.48	0.86
	$K=25$	56.52	63.41	50.98	90.00	0.73		$K=25$	18.53	10.76	66.67	90.27	0.85
	$K=30$	17.03	9.76	66.67	90.25	0.69		$K=30$	26.53	17.17	58.33	89.75	0.86
	$K=35$	17.03	9.76	66.67	90.25	0.69		$K=35$	23.52	14.63	60.00	89.75	0.87
	$K=40$	20.84	12.20	71.43	90.50	0.71		$K=40$	23.30	14.63	57.14	90.48	0.85

TABLE 16. Evaluation GNN model in Scenario 7 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K=5$	88.01	88.37	87.66	88.86
$K=10$	88.63	89.84	87.46	90.15
$K=15$	87.11	88.42	85.84	88.63
$K=20$	87.43	88.40	86.49	88.87
$K=25$	86.68	88.30	85.11	88.40
$K=30$	89.47	90.09	88.85	**91.51**
$K=35$	88.95	89.61	88.29	90.11
$K=40$	89.09	89.83	88.36	90.25

Table. 19 illustrates the implementation results of machine learning (ML) models, with DT achieving the highest accuracy at 90.75%. However, upon closer examination of F1-scores, SVM with a graph generated using K = 40 demonstrates optimal performance, boasting an F1-score of 69.75%. Notably, the DT model applied to the dataset generated with K = 15 achieves the highest F1-score among the accuracy-leading models, suggesting it as a preferable choice over LR for assessing credit risk.

Analyzing ROC curves across different K values, MLP consistently outperforms other models at K = 5, 10, 15, 25, 30, 35, and 40, with AUC values ranging from 0.86 to 0.88. Additionally, LR and MLP demonstrate dominance at K = 20, both achieving an AUC value of 0.86. These results indicate the superior positioning of MLP within the ROC space compared to alternative models, suggesting its effectiveness in credit risk assessment.

Table. 20 reveals a notable consistency in performance metrics, including "F1-score", "Recall", "Precision", and "Accuracy", which are observed to fall within a narrow range, spanning from 83.43 percent to 92.11 percent. Such stability reflects a marked difference from conventional machine learning model implementations, highlighting a distinctive trait of the GNN model. Notably, the assessment of credit risk attains its highest accuracy at K = 10, reaching 92.11 percent.

D. SUMMARY OF RESULTS AND DISCUSSION

A comparative analysis was conducted in this study to evaluate the performance of four ML models: (i) decision tree, (ii) logistic regression, (iii) support vector machine and (iv) multi-layer perceptron in comparison to the GNN model. This comparison was carried out across nineteen distinct feature sets derived from financial data of 2000 SMEs operating within the energy sector supply chain in 2022. The datasets consisted of 93 different financial ratios and were analyzed under nine distinct scenarios:

Scenario 1: Models are applied to a feature set containing all financial ratios of companies. This approach offers a comprehensive overview of the dataset but may suffer from redundancy and noise in the features.

- **Hypothesis 1:** The initial features (Basic Feature Set) are insufficient for accurate credit risk prediction.

TABLE 17. Evaluation of machine learning models in Scenario 8 (percent).

	K-Values	F1-score	Recall	Precision	Accuracy	AUC		K-Values	F1-score	Recall	Precision	Accuracy	AUC
Logistic Regression	$K=5$	35.29	21.95	90.00	**91.75**	0.76	Multi-layer Perceptron	$K=5$	38.46	24.39	90.91	**92.00**	0.75
	$K=10$	32.00	19.51	88.89	91.50	0.80		$K=10$	37.04	24.39	76.92	91.50	0.80
	$K=15$	29.17	17.07	100.00	91.50	0.80		$K=15$	32.65	19.51	100.00	91.75	0.81
	$K=20$	32.00	19.51	88.89	91.50	0.84		$K=20$	32.00	19.51	88.89	91.50	0.84
	$K=25$	32.73	21.95	64.29	90.75	0.83		$K=25$	32.73	21.95	64.29	90.75	0.83
	$K=30$	23.08	14.63	54.55	90.00	0.86		$K=30$	23.08	14.63	54.55	90.00	0.84
	$K=35$	27.45	17.07	70.00	90.75	0.85		$K=35$	41.98	41.46	42.50	88.25	0.84
	$K=40$	30.19	19.51	66.67	90.75	0.84		$K=40$	29.63	19.51	61.54	90.50	0.83
Decision Trees	$K=5$	35.29	21.95	90.00	**91.75**	0.69	Support Vector Machines	$K=5$	17.78	9.76	100.00	90.75	0.59
	$K=10$	47.31	53.66	42.31	87.75	0.68		$K=10$	25.53	14.63	100.00	91.25	0.73
	$K=15$	46.51	48.78	44.44	88.50	0.65		$K=15$	29.17	17.07	100.00	**91.50**	0.68
	$K=20$	42.11	48.78	37.04	86.25	0.71		$K=20$	32.00	19.51	88.89	**91.50**	0.64
	$K=25$	40.48	41.46	39.53	87.50	0.66		$K=25$	32.73	21.95	64.29	90.75	0.65
	$K=30$	45.98	48.78	43.48	88.25	0.72		$K=30$	37.29	26.83	61.11	90.75	0.69
	$K=35$	27.45	17.07	70.00	90.75	0.73		$K=35$	27.45	17.07	70.00	90.75	0.72
	$K=40$	48.72	46.34	51.35	90.00	0.76		$K=40$	21.74	12.20	100.00	91.00	0.72

TABLE 18. Evaluation GNN model in Scenario 8 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K=5$	88.15	87.86	88.44	88.23
$K=10$	90.80	90.92	90.68	91.40
$K=15$	90.14	90.22	90.07	90.59
$K=20$	90.36	90.79	89.94	91.17
$K=25$	91.04	91.36	90.72	91.76
$K=30$	91.91	92.32	91.51	**92.46**
$K=35$	90.51	90.82	90.20	91.31
$K=40$	91.19	91.60	90.79	91.76

- **Results:** As expected, the GNN model outperforms ML models, but the accuracy is lower compared to scenarios involving feature selection or network-based features.

Scenario 2: Models are implemented on a feature set comprising selected ratios obtained using the BallMapper algorithm. This method focuses on extracting relevant features based on the distribution of data points in multidimensional space.

- **Hypothesis 2:** BallMapper-selected features enhance the model's predictive performance by reducing noise and redundancy.
- **Results:** The GNN model achieves higher accuracy and F1-score compared to Scenario 1, confirming that BallMapper-selected features are more effective for credit risk prediction.

Scenario 3: Models are applied to a feature set containing selected ratios derived from intercompany correlation relationships. This approach emphasizes the correlation between different financial indicators as a basis for feature selection.

- **Hypothesis 3:** Correlation-based feature selection provides a simpler yet effective method for feature extraction, though it may be less effective than BallMapper in modeling non-linear relationships.
- **Results:** While ML and GNN models perform better than Scenario 1, the performance of GNN in this scenario is slightly inferior to Scenario 2, validating the advantage of BallMapper in capturing non-linear relationships.

Scenarios 4 and 5: Models are implemented on feature sets containing network features constructed based on either the BallMapper algorithm or intercompany correlation relationships, respectively. These scenarios leverage graph-based representations of the data to capture complex relationships and structures within the dataset.

- **Hypothesis 2:** Network-based features better model interdependencies and systemic risks in the energy supply chain.
- **Results:** The GNN model demonstrates significant improvements in accuracy and F1-score compared to Scenarios 1–3, with Scenario 4 showing superior results, highlighting the effectiveness of graph-based features derived from BallMapper.

Scenario 6: Combines the total of all financial ratios with network features constructed using the BallMapper algorithm.

TABLE 19. Evaluation of machine learning models in Scenario 9 (percent).

	K-Values	F1-score	Recall	Precision	Accuracy	AUC		K-Values	F1-score	Recall	Precision	Accuracy	AUC
Logistic Regression	$K=5$	13.04	7.32	60.00	90.00	0.83	Multi-layer Perceptron	$K=5$	25.45	17.07	50.00	89.75	0.87
	$K=10$	13.04	7.32	60.00	90.00	0.82		$K=10$	27.59	19.51	47.06	89.50	0.87
	$K=15$	16.67	9.76	57.14	90.00	0.83		$K=15$	22.22	14.63	46.15	89.50	0.86
	$K=20$	13.04	7.32	60.00	90.00	0.86		$K=20$	21.82	14.63	42.86	89.25	0.86
	$K=25$	17.78	9.76	100.00	**90.75**	0.85		$K=25$	27.59	19.51	47.06	89.50	0.88
	$K=30$	17.39	9.76	80.00	90.50	0.83		$K=30$	21.82	14.63	42.86	89.25	0.87
	$K=35$	17.02	9.76	66.67	90.25	0.82		$K=35$	24.56	17.07	43.75	89.25	0.86
	$K=40$	17.78	9.76	100.00	**90.75**	0.86		$K=40$	20.83	12.20	71.43	**90.50**	0.88
Decision Trees	$K=5$	51.16	53.66	48.89	89.50	0.76	Support Vector Machines	$K=5$	23.06	15.07	49.05	89.75	0.86
	$K=10$	51.16	53.66	48.89	89.50	0.75		$K=10$	23.08	14.63	54.55	89.50	0.86
	$K=15$	53.16	51.22	55.26	**90.75**	0.70		$K=15$	55.17	58.54	52.17	88.50	0.82
	$K=20$	53.49	56.10	51.11	90.00	0.75		$K=20$	29.63	19.51	61.54	59.50	0.85
	$K=25$	53.93	58.54	50.00	89.75	0.75		$K=25$	22.64	14.63	50.00	90.25	0.84
	$K=30$	49.41	51.22	47.73	89.25	0.76		$K=30$	29.09	19.51	57.14	**90.50**	0.85
	$K=35$	51.85	51.22	52.50	90.25	0.76		$K=35$	41.98	41.46	42.50	90.25	0.77
	$K=40$	51.16	53.66	48.89	89.50	0.78		$K=40$	69.77	73.17	66.67	90.25	0.85

TABLE 20. Evaluation GNN model in Scenario 9 (percent).

K-Values	F1-score	Recall	Precision	Accuracy
$K=5$	88.15	87.86	88.44	88.23
$K=10$	90.80	90.92	90.68	91.40
$K=15$	90.14	90.22	90.07	90.59
$K=20$	90.36	90.79	89.94	91.17
$K=25$	91.04	91.36	90.72	91.76
$K=30$	91.91	92.32	91.51	**92.46**
$K=35$	90.51	90.82	90.20	91.31
$K=40$	91.19	91.60	90.79	91.76

This integrated approach aims to leverage both numerical and graph-based representations of the data for improved predictive performance.

- **Hypothesis 4:** Combining BallMapper-selected features with graph-based features yields the best performance in credit risk assessment.
- **Results:** The highest accuracy of 93.56% and superior F1-score for the GNN model are achieved in this scenario, confirming the hypothesis that integrating numerical and graph-based features leads to the most robust credit risk prediction.

Scenario 7: Models are applied to a dataset comprising the total of all financial ratios and network features constructed based on intercompany correlation relationships.

- **Hypothesis 4:** Integration of financial ratios with correlation-based graph features improves prediction accuracy, though it may be less effective than Scenario 6.
- **Results:** While GNN achieves high accuracy and F1-score, the performance is slightly below Scenario 6, reinforcing the greater efficacy of BallMapper-based features.

Scenarios 8 and 9: Involve implementing models on feature sets combining selected ratios obtained using either the BallMapper algorithm or intercompany correlation relationships with constructed network features. These scenarios aim to assess the combined impact of feature selection methods and graph-based representations on model performance.

- **Hypothesis 4:** The combination of BallMapper features and network features provides a balanced trade-off between precision and recall, ensuring optimal prediction outcomes.
- **Results:** Scenario 8 outperforms Scenario 9, further confirming the strength of BallMapper-selected features and their integration with graph-based data.

As illustrated in Fig. 10(a), across all investigated scenarios, the GNN model achieved the highest accuracy in evaluating the credit risk of companies through predicting their likelihood of bankruptcy over the next two to three years. This underscores the superior predictive capability of GNN in comparison to other models examined in our study.

The highest accuracy achieved in implementing the GNN model stands at 93.56%, obtained in Scenario 6. Overall, the performance of the GNN surpasses that of ML

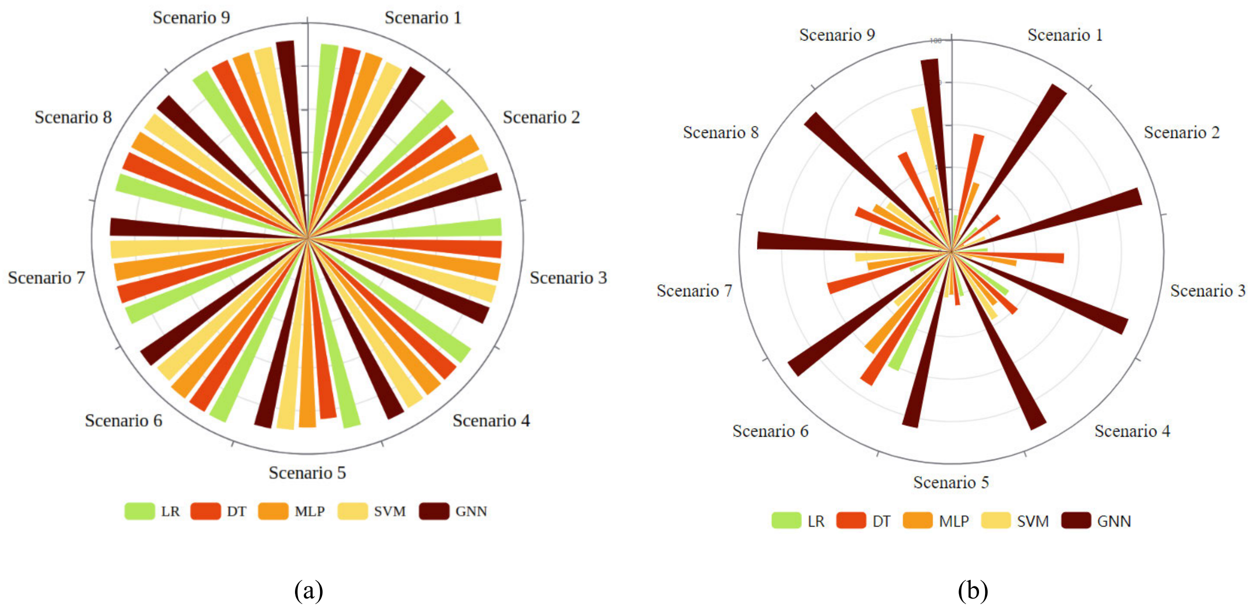


FIGURE 10. **(a) Comparison of each scenarios according to the accuracy of the implemented models (b) Comparison of each of the scenarios based on the "F1-score" of the implemented models.**

models. However, it is noteworthy that the most effective approach in evaluating credit risk in the studied research lies in incorporating all financial ratios of companies along with graph-based features constructed using the BallMapper algorithm. This indicates the importance of leveraging both financial metrics and graph-based features for robust credit risk assessment.

Furthermore, as depicted in Fig. 10(b), the F1-score in the implementation of the GNN model notably outperforms that of traditional machine learning models. The notable contrast highlights the GNN model's superior performance in attaining a delicate equilibrium between precision and recall, which is pivotal for effective credit risk assessment. Since the F-score is derived from recall and precision, the graphs related to these two metrics exhibit similar behavior to Fig. 10(b). Therefore, we have limited our discussion to the graph related to the F-score.

It's worth noting that in Scenario 6, the discrepancy in "F1-score" between the implementation of ML models and the GNN model is less significant. This indicates that the combination of graph-based features constructed using the BallMapper algorithm alongside other financial ratios yields more favorable results in evaluating the credit risk of companies. This underscores the significance of integrating graph-based features alongside financial ratios for more effective credit risk assessment.

To gain insight into the crucial features influencing our model, we can generate a visualization depicting SHAP (Shapley Additive explanations) values for each feature across all samples. Fig. 11 organizes features based on the cumulative sum of SHAP value magnitudes throughout the feature sets, utilizing SHAP values to illustrate the distribution of feature's Influence on the BM-GNN's results. Additionally, feature values are represented by a color scale, with higher values depicted in red and lower values in blue.

Based on Fig. 11, it is clear that certain features play a pivotal role in predicting bankruptcy. Notably, "Operating Funds to Liability", "Net Value Per Share (Before Interest and Depreciation After Tax)", "Cash Flow to Total Assets", and "Cash Reinvestment" emerge as the most influential factors. These features exhibit a positive impact on the likelihood of bankruptcy, implying that as their values increase, so does the probability of bankruptcy. Conversely, "Net Profit Growth Rate" exerts a negative influence on bankruptcy probability; as this feature's value rises, the likelihood of bankruptcy decreases.

In contemporary management practices, it's commonplace for decision makers to seek comprehensive insights beyond mere identification of key independent variables and forecast performance. Their quest extends to understanding the nuanced impacts of these variables on anticipated outcomes. This holistic approach empowers financial institution (FI) managers to mitigate financing credit risks, aids SMEs managers in enhancing their financing capabilities, and equips corporate entity (CE) managers to navigate and minimize joint liability credit risks.

Pioneered by Friedman in 2001, the methodology of Partial Dependency Plot (PDP) analysis emerges as a potent tool for dissecting the influence of independent variables on predicted responses through graphical representation. PDP excels in unveiling both linear and nonlinear relationships existing amidst independent variables and their corresponding predictive outcomes, leveraging various regression models such as regression trees. Its ability to generate line plots depicting predicted responses against individual features, while factoring in the collective influence of other independent variables, renders it a versatile analytical asset [51].

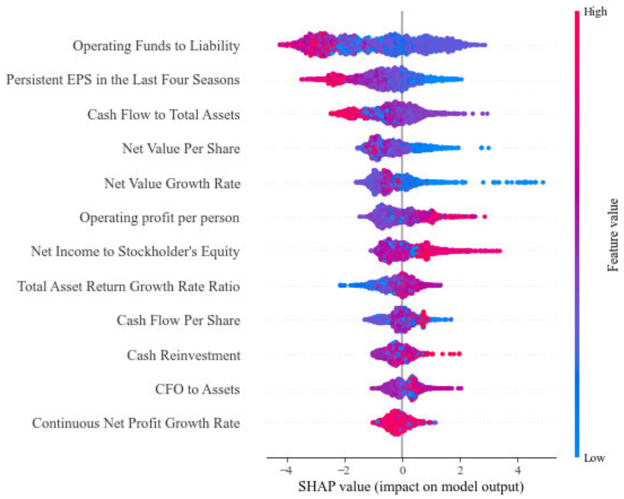


FIGURE 11. **The influence between indicators and results.**

In the forthcoming discourse, we embark on a PDP based exploration to dissect the impact of variable v_i on the likelihood of non-risky outcomes for SMEs. This section delves into a nuanced examination of each variable's role in shaping risk assessment, mirroring the insights encapsulated in Fig. 12 of our analysis.

Fig. 12(a) illustrates the correlation between an upward trend in the net value per share index and a decreased probability of company bankruptcies for SMEs in SCF. This index represents the difference between a company's assets and liabilities divided by its outstanding shares. A higher net value per share indicates a robust financial stance, suggesting lower credit risk, while a lower value may signal heightened financial vulnerability. Therefore, the net value per share serves as a critical determinant in forecasting credit risk assessment in SCF, offering indispensable guidance for evaluating creditworthiness and implementing risk mitigation strategies. Fig. 12(b) shows the persistent Earnings Per Share (EPS) over the last four seasons for SMEs in SCF, indicating their financial stability and performance over time. An increase in this index is linked with a decrease in the likelihood of company bankruptcies. Fig. 12(c) depicts the cash flow per share (CFPS), which reflects operational efficiency and the ability to generate cash from core activities, crucial for sustainability. A higher CFPS indicates financial stability, reducing risk perception, while weak CFPS raises concerns about financial stability and ability to repay debts. The likelihood of company bankruptcies decreases with an increase in CFPS up to approximately 3% but escalates beyond this threshold, reaching 20%. Fig. 12(d) highlights the importance of continuous improvement in the net profit growth rate for SMEs in SCF. It signifies enhanced profitability, effective operational management, and potential market share expansion. In credit risk assessment, a positive net profit growth rate indicates financial stability and operational efficiency, making SMEs more favorable to financial service providers. Conversely, a decreasing or negative net profit rate may signal financial troubles or operational inefficiencies, increasing credit risk. Continuous monitoring of the net profit growth rate is crucial in evaluating SME creditworthiness in SCF, providing insights into their financial health and future prospects. The figure shows that as the net profit growth rate increases up to 0.2, SMEs in the energy sector supply chain have a 21% probability of bankruptcy, which decreases significantly beyond this threshold. Fig. 12(e) presents the net value growth rate, a critical indicator for SMEs in SCF, reflecting their financial health and trajectory. As depicted, higher values of this indicator correspond to a probability of bankruptcy for companies falling below 10%. A positive net value growth rate suggests successful business operations, expansion, or efficiency improvements within the supply chain, thereby lowering perceived credit risk. Conversely, a negative or stagnant growth rate may raise concerns about the SME's financial capacity, potentially elevating credit risk. Financial institutions often use the net value growth rate in risk assessment models to inform financing decisions, with higher growth rates associated with lower credit risk. Monitoring and analyzing the net value growth rate are essential for evaluating SMEs' creditworthiness and financial stability in SCF. Fig. 12(f) illustrates the total asset return growth rate for SMEs in SCF, indicating the rate at which the value of assets owned by these businesses is increasing over time. As depicted, as the total asset return growth rate ratio increases up to 0.25, companies face a probability of approximately 22% of bankruptcy. However, beyond this threshold, the probability diminishes below 10%. Yet, considering various economic and environmental factors, the likelihood of bankruptcy increases again, reaching 15%. This growth rate is a crucial indicator of SMEs' performance and financial health within the SCF context. A higher total asset return growth rate suggests significant expansion and healthy investment returns, implying lower credit risk. Conversely, a negative or declining total asset return growth rate may indicate financial instability and higher credit risk. Hence, financial institutions and lenders should consider the total asset return growth rate when evaluating the creditworthiness of SMEs in SCF. Fig. 12(g) highlights cash reinvestment, which involves reinvesting cash generated from operational activities into the supply chain to optimize liquidity and efficiency. An increase in cash reinvestment by up to 30% predicts a probability of bankruptcy for companies exceeding 16%. However, beyond this threshold, the probability fluctuates between 5% to 16%. Ultimately, with a cash reinvestment value from 40% to 100%, companies' bankruptcy probability reaches approximately 11%. This strategic approach empowers SMEs to enhance working capital management, foster stronger supplier relationships, and bolster overall supply chain resilience. Efficient cash reinvestment mitigates liquidity risks, ensures timely payment to suppliers, and

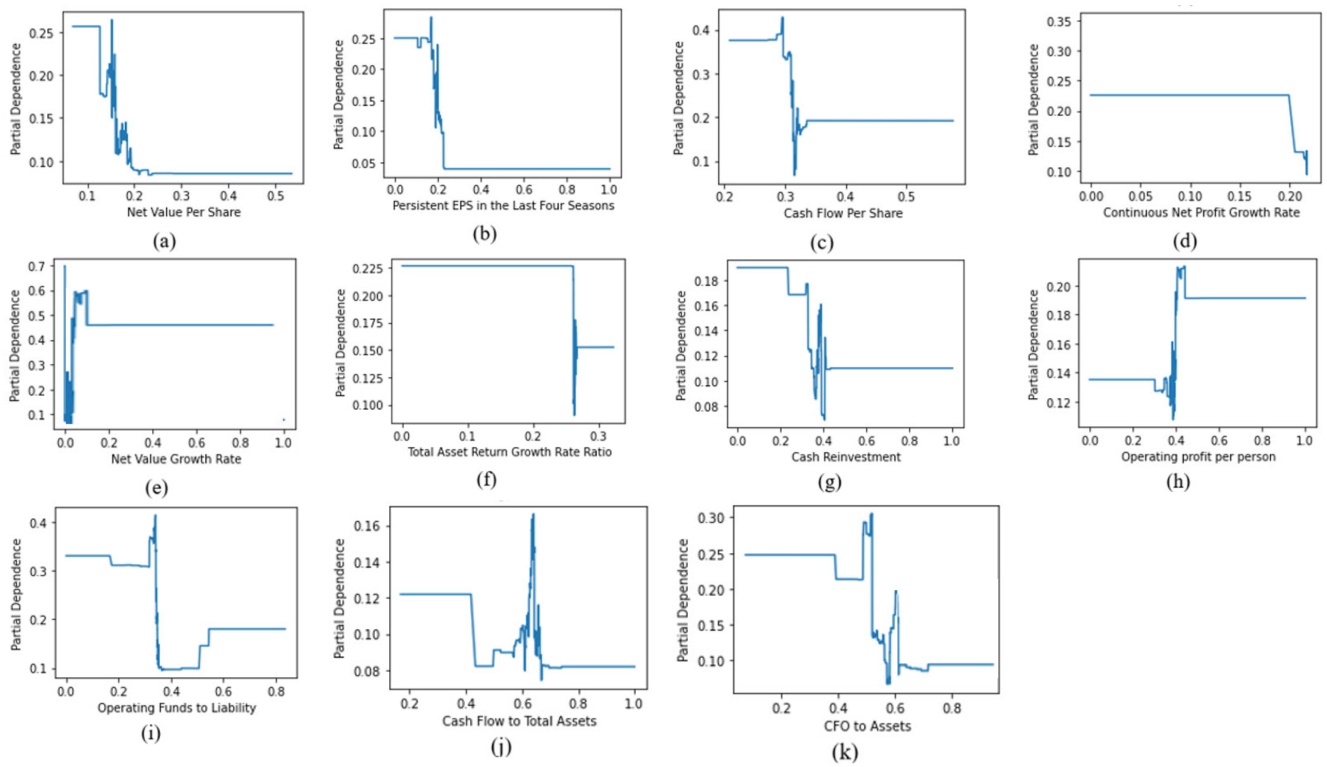


FIGURE 12. The line PDP of the top features of GNN.

demonstrates financial stability, which are crucial factors in assessing creditworthiness. Incorporating cash reinvestment insights into credit risk assessment models can enhance predictive accuracy and comprehensively evaluate SMEs' creditworthiness in SCF contexts. Fig. 12(h) underscores the significance of operating profit per person, a vital metric for SMEs providing insights into business efficiency and profitability on a per-employee basis. As operating profit per person increases up to 0.4, the probability of bankruptcy within recent years reaches approximately 14%. Beyond this threshold, under various environmental and economic conditions, this probability exceeds 20%. However, once the indicator surpasses 0.42, the probability stabilizes at 19%. This metric reflects how efficiently the company utilizes labor resources to generate profits. Higher operating profit per person suggests operational efficiency and better financial health, potentially lowering credit risk. Conversely, lower operating profit per person may indicate financial challenges or inefficiencies, raising credit risk. Monitoring this metric is instrumental in assessing SME creditworthiness in SCF and guiding risk management strategies for financial institutions. Fig. 12(i) highlights operating funds liability, representing the debts incurred by businesses in their day-to-day operations within the supply chain. As the operating funds to liability ratio increases up to 0.35, the probability of bankruptcy rises from 30% to over 40%. Subsequently, the probability sharply decreases to below 10%. However, surpassing the 0.5 threshold increases the likelihood of bankruptcy to 15%. Operating funds liability reflects the SME's financial health and liquidity management. Higher liability suggests potential liquidity challenges or cash flow inefficiencies, leading to increased credit risk. Managing operating funds liability is crucial for accurate credit risk assessment and ensuring SMEs' financial stability in SCF. Fig. 12(j) depicts the cash flow to total assets ratio, showcasing how efficiently an SME utilizes its assets to generate cash flow. As this ratio increases, indicating improved liquidity, the probability of bankruptcy decreases. It offers valuable insights into the SME's ability to convert assets into cash, crucial for operational continuity and financial commitments. A higher ratio suggests effective asset management and lower credit risk, while a lower ratio may indicate inefficiencies and higher credit risk. This ratio serves as a significant indicator in assessing credit risk, reflecting the SME's ability to generate cash flow from its assets to meet financial commitments. Fig. 12(k) shows the CFO to asset ratio, indicating the efficiency of cash flow management relative to the total assets of the companies. An increase in this ratio correlates with a decrease in the probability of bankruptcy for companies, demonstrating improved financial health and operational efficiency. A higher CFO to asset ratio suggests efficient asset utilization and strong cash flow generation, positively impacting credit risk assessment. Conversely, a lower ratio may raise concerns about liquidity and financial resilience, potentially increasing perceived credit risk. Therefore, analyzing the CFO to asset ratio is crucial for accurately assessing the credit risk in SCF, providing valuable insights into their financial performance and resilience. Fig. 12(l) illustrates the net income to stockholders' equity ratio, indicating a company's profitability relative to the equity invested by shareholders. An increase in this ratio up to 0.8 is associated with a probability of bankruptcy around 22%, decreasing to approximately 10% beyond that threshold. This ratio serves as a measure of financial performance and efficiency in generating shareholder returns for SMEs in the SCF field. A higher ratio suggests effective utilization of equity to generate profits, indicating lower credit risk. Conversely, a lower ratio may indicate financial distress, potentially elevating credit risk. Therefore, this ratio plays a crucial role in assessing credit risk in SCF, providing insights into the SME's ability to generate profits relative to shareholder equity and aiding financial institutions in making informed predictions about creditworthiness and financial stability.

1) FEATURE IMPORTANCE ANALYSIS AND SECTOR-SPECIFIC PATTERNS

The energy sector's selection indeed distinguishes our study, as SMEs in this sector face unique financial challenges and risk factors. To explore sector-specific patterns, we conducted a detailed analysis of the feature importance results, focusing on how certain financial indicators uniquely impact credit risk for SMEs in the energy sector.

2) KEY FINDINGS

- **Operating Funds to Liability:** This feature emerged as a critical indicator of financial health in the energy sector. SMEs with higher ratios of operating funds to liabilities were found to have a significantly lower risk of bankruptcy. This highlights the importance of liquidity and efficient fund management in mitigating financial distress in an industry characterized by high capital intensity and long operational cycles.
- **Net Value Per Share (Before Interest and Depreciation After Tax):** Our analysis revealed that SMEs with higher net values per share tended to be more resilient to financial shocks. This suggests that maintaining a strong equity base and effective asset management are crucial for sustaining operations in the volatile energy market.
- **Cash Flow:** Consistent positive cash flow was identified as a vital factor in reducing credit risk. Energy sector SMEs that managed to maintain stable cash flows despite market fluctuations were better positioned to meet their financial obligations, indicating the importance of robust cash flow management practices.
- **Inventory Turnover Ratio:** A higher inventory turnover ratio was associated with lower credit risk, reflecting efficient inventory management and quicker conversion of inventory into sales. This is particularly important in the energy sector, where inventory can include costly and perishable items.
- **Receivables Turnover Ratio:** A higher receivables turnover ratio indicated better credit control and faster collection of receivables. SMEs that excelled in managing their accounts receivable were less likely to experience liquidity issues, underscoring the significance of effective credit policies and debtor management.

3) SECTOR-SPECIFIC PATTERNS

- **Volatility Management:** The energy sector is subject to significant price and demand volatility. SMEs that exhibited strong financial indicators related to liquidity and asset management were better equipped to handle such volatility, reducing their overall credit risk.
- **Capital Intensity:** The high capital requirements in the energy sector mean that SMEs must manage their resources efficiently. Features related to capital utilization, such as net value per share and operating funds to liability, were crucial in differentiating low-risk SMEs from their high-risk counterparts.
- **Supply Chain Dependencies:** The interconnected nature of the energy supply chain means that disruptions can have cascading effects. SMEs with robust inventory and receivables management were able to mitigate these risks more effectively.

4) STRENGTHENING THE ARGUMENT FOR CONTRIBUTION

By identifying and analyzing these sector-specific patterns, our research highlights the unique financial dynamics that impact credit risk in the energy sector. This deeper exploration not only strengthens the predictive capability of our BM-GNN model but also underscores its practical value for stakeholders in the energy sector. Financial institutions and SMEs can leverage these insights to enhance their credit risk assessment frameworks and implement targeted risk mitigation strategies.

In summary, our expanded analysis in Section D demonstrates that the BM-GNN model effectively captures critical financial indicators specific to the energy sector, providing a comprehensive tool for predicting SME credit risk. This contributes to the broader understanding of credit risk management in high-stakes, capital-intensive industries like energy.

V. CONCLUSION

Assessing credit risk has been both challenging and pivotal for financial institutions across various societies. Thus far, most studies have been conducted based on independent variables related to companies' demographic, psychographic, and financial characteristics. In this study, however, we have taken a different approach from previous research and utilized the BallMapper algorithm to isolate the financial indicators that have the most significant impact on predicting companies' credit risk. Subsequently, by forming an adjacency matrix, we integrated network features with other attributes in various ways and evaluated the performance of four ML models alongside a GNN model. We aimed to innovate in our approach to credit risk assessment by leveraging advanced techniques such as the BallMapper algorithm and integrating network features with traditional predictors. This novel methodology allows us to potentially uncover deeper insights into the drivers of credit risk and enhance predictive accuracy in financial decision-making processes. This study proposes a multi-modal approach, termed BM-GNN, aimed at enhancing credit risk prediction for SMEs. By integrating BallMapper with GNN analysis techniques, this approach offers a comprehensive and sophisticated framework for risk management, surpassing the capabilities of available models highlighted in previous research. Leveraging the topological properties of data, topological data analysis provides deeper insights into credit risk factors, thereby optimizing the accuracy and reliability of credit risk predictions for SMEs. In conclusion, the integration of GNNs and topological data analysis techniques (BallMapper in this study) offers a promising avenue for enhancing credit risk prediction capabilities, particularly within the context of SMEs and SCF. By leveraging the network relationships and non-financial data sources provided by GNNs, alongside the analytical depth of topological data analysis, a more comprehensive evaluation of credit risk for SMEs engaged in supply chain finance can be achieved. This multi-modal approach, exemplified by the BM-GNN model, not only improves the reliability and accuracy of credit risk evaluation for SMEs but also enables financial institutions to better anticipate default probabilities and manage credit risk effectively within the dynamic landscape of SCF. Hence, the integration of GNNs and topological data analysis signifies a substantial progression in credit risk prediction, equipping institutions with deeper insights and enhanced decision-making capabilities to support the financial requirements of SMEs within supply chain ecosystems.

REFERENCES

[1] Y. Yang, X. Chen, J. Gu, and H. Fujita, "Alleviating financing constraints of SMEs through supply chain," *Sustainability*, vol. 11, no. 3, p. 673, Jan. 2019.

[2] Z. Wang and G. Liang, "Game analysis on supply chain inventory financing under demand uncertainties," in *Proc. 2nd Int. Conf. Data Sci. Bus. Anal. (ICDSBA)*, Sep. 2018, pp. 441–447.

[3] Z. Fatima, V. Oksman, and R. Lahdelma, "Enabling small medium enterprises (SMEs) to become leaders in energy efficiency using a continuous maturity matrix," *Sustainability*, vol. 13, no. 18, p. 10108, Sep. 2021.

[4] F. Manzoor, L. Wei, and N. Sahito, "The role of SMEs in rural development: Access of SMEs to finance as a mediator," *PLoS ONE*, vol. 16, no. 3, Mar. 2021, Art. no. e0247598.

[5] Y. Zhu, L. Zhou, C. Xie, G.-J. Wang, and T. V. Nguyen, "Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach," *Int. J. Prod. Econ.*, vol. 211, pp. 22–33, May 2019, doi: 10.1016/j.ijpe.2019.01.032.

[6] L. Kong, G. Zheng, and A. Brintrup, "A federated machine learning approach for order-level risk prediction in supply chain financing," *Int. J. Prod. Econ.*, vol. 268, Feb. 2024, Art. no. 109095, doi: 10.1016/j.ijpe.2023.109095.

[7] C. Wu, J. Liu, and H. Zhang, "Data ecology and accurate portrait: Optimization of credit risk system for SMEs in supply chain finance based on big data technology," *J. Risk Anal. Crisis Response*, vol. 11, no. 4, Jan. 2022.

[8] M. Gidea and Y. Katz, "Topological data analysis of financial time series: Landscapes of crashes," *Phys. A, Stat. Mech. Appl.*, vol. 491, pp. 820–834, Feb. 2018.

[9] B. Liu, I. Li, J. Yao, Y. Chen, G. Huang, and J. Wang, "Unveiling the potential of graph neural networks in SME credit risk assessment," 2024, *arXiv:2409.17909*.

[10] C. Huang, F. T. S. Chan, and S. H. Chung, "Recent contributions to supply chain finance: Towards a theoretical and practical research agenda," *Int. J. Prod. Res.*, vol. 60, no. 2, pp. 493–516, Jan. 2022.

[11] S. Saberi, M. Kouhizadeh, J. Sarkis, and L. Shen, "Blockchain technology and its relationships to sustainable supply chain management," *Int. J. Prod. Res.*, vol. 57, no. 7, pp. 2117–2135, Apr. 2019, doi: 10.1080/00207543.2018.1533261.

[12] F. Olan, S. Liu, J. Suklan, U. Jayawickrama, and E. O. Arakpogun, "The role of artificial intelligence networks in sustainable supply chain finance for foodand drink industry," *Int. J. Prod. Res.*, vol. 60, no. 14, pp. 4418–4433, Jul. 2022.

[13] J. Chen, A. S. Sohal, and D. I. Prajogo, "Supply chain operational risk mitigation: A collaborative approach," *Int. J. Prod. Res.*, vol. 51, no. 7, pp. 2186–2199, Apr. 2013.

[14] A. Ghadge, S. K. Jena, S. Kamble, D. Misra, and M. K. Tiwari, "Impact of financial risk on supply chains: A manufacturer-supplier relational perspective," *Int. J. Prod. Res.*, vol. 59, no. 23, pp. 7090–7105, Dec. 2021.

[15] P. Dutta, T.-M. Choi, S. Somani, and R. Butala, "Blockchain technology in supply chain operations: Applications, challenges and research opportunities," *Transp. Res. E, Logistics Transp. Rev.*, vol. 142, Oct. 2020, Art. no. 102067.

[16] Z. Li, X. Liang, Q. Wen, and E. Wan, "The analysis of financial network transaction risk control based on blockchain and edge computing technology," *IEEE Trans. Eng. Manag.*, vol. 71, pp. 5669–5690, 2024, doi: 10.1109/TEM.2024.3364832.

[17] G. Bressanelli, M. Perona, and N. Saccani, "Challenges in supply chain redesign for the circular economy: A literature review and a multiple case study," *Int. J. Prod. Res.*, vol. 57, no. 23, pp. 7395–7422, Dec. 2019.

[18] T.-M. Choi, A. A. Taleizadeh, and X. Yue, "Game theory applications in production research in the sharing and circular economy era," *Int. J. Prod. Res.*, vol. 58, no. 1, pp. 118–127, 2020.

[19] Y.-S. Huang, R.-S. Ho, and C.-C. Fang, "Quantity discount coordination for allocation of purchase orders in supply chains with multiple suppliers," *Int. J. Prod. Res.*, vol. 53, no. 22, pp. 6653–6671, Nov. 2015.

[20] D. A. Phan, T. L. H. Vo, and A. N. Lai, "Supply chain coordination under trade credit and retailer effort," *Int. J. Prod. Res.*, vol. 57, no. 9, pp. 2642–2655, May 2019.

[21] M. Dora, A. Kumar, S. K. Mangla, A. Pant, and M. M. Kamal, "Critical success factors influencing artificial intelligence adoption in food supply chains," *Int. J. Prod. Res.*, vol. 60, no. 14, pp. 4621–4640, Jul. 2022.

[22] A. Gunasekaran, N. Subramanian, and S. Rahman, "Supply chain resilience: Role of complexities and strategies," *Int. J. Prod. Res.*, vol. 53, no. 22, pp. 6809–6819, Nov. 2015.

[23] D. Ivanov, B. Sokolov, I. Solovyeva, A. Dolgui, and F. Jie, "Dynamic recovery policies for time-critical supply chains under conditions of ripple effect," *Int. J. Prod. Res.*, vol. 54, no. 23, pp. 7245–7258, Dec. 2016.

[24] P. Wichmann, A. Brintrup, S. Baker, P. Woodall, and D. McFarlane, "Extracting supply chain maps from news articles using deep neural networks," *Int. J. Prod. Res.*, vol. 58, no. 17, pp. 5320–5336, Sep. 2020.

[25] G. Baryannis, S. Validi, S. Dani, and G. Antoniou, "Supply chain risk management and artificial intelligence: State of the art and future research directions," *Int. J. Prod. Res.*, vol. 57, no. 7, pp. 2179–2202, Apr. 2019, doi: 10.1080/00207543.2018.1530476.

[26] M. Kamalahmadi and M. Mellat-Parast, "Developing a resilient supply chain through supplier flexibility and reliability assessment," *Int. J. Prod. Res.*, vol. 54, no. 1, pp. 302–321, Jan. 2016.

[27] B. Nepal and O. P. Yadav, "Bayesian belief network-based framework for sourcing risk analysis during supplier selection," *Int. J. Prod. Res.*, vol. 53, no. 20, pp. 6114–6135, Oct. 2015.

[28] R. Wang, C. Yu, and J. Wang, "Construction of supply chain financial risk management mode based on Internet of Things," *IEEE Access*, vol. 7, pp. 110323–110332, 2019.

[29] Y. Wang, "Research on supply chain financial risk assessment based on blockchain and fuzzy neural networks," *Wireless Commun. Mobile Comput.*, vol. 2021, no. 1, Jan. 2021, Art. no. 5565980.

[30] W. A. Abbasi, Z. Wang, Y. Zhou, and S. Hassan, "Research on measurement of supply chain finance credit risk based on Internet of Things," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 9, Sep. 2019, Art. no. 155014771987400.

[31] N. B. Kalyan, K. Ahmad, F. Rahi, C. Shelke, and S. M. Basha, "Application of Internet of Things and machine learning in improving supply chain financial risk management system," in *Proc. IEEE 2nd Int. Conf. Ind. Electron., Develop. Appl. (ICIDeA)*, Sep. 2023, pp. 211–216, doi: 10.1109/icidea59866.2023.10295182.
[32] Y. Fu, M. Li, X. Sun, and B. Shang, "A prediction model of credit risk in SCF based on IG-GA-SVM," in *Proc. 2nd Int. Conf. Adv. Algorithms Signal Image Process. (AASIP)*. Bellingham, WA, USA: SPIE, Nov. 2022, p. 28.
[33] L. Zhu, W. Liu, R. Zhang, and B. Dong, "Credit risk evaluation of supply chain finance based on K-means-SVM model," in *Proc. 4th Int. Conf. Appl. Mach. Learn. (ICAML)*, Jul. 2022, pp. 410–413.
[34] W. Chen, Z. Li, and Z. Xiao, "On credit risk contagion of supply chain finance under COVID-19," *J. Math.*, vol. 2021, pp. 1–13, Sep. 2021.
[35] Q. Zhou, L. Wang, L. Juan, S. Zhou, and L. Li, "The study on credit risk warning of regional listed companies in China based on logistic model," *Discrete Dyn. Nature Soc.*, vol. 2021, pp. 1–8, May 2021.
[36] Y. Mao, S. Liu, and D. Gong, "A hybrid technological innovation text mining, ensemble learning and risk scorecard approach for enterprise credit risk assessment," *Tehnički vjesnik*, vol. 30, no. 6, pp. 1692–1703, 2023.
[37] Y. Wei and H. Hasan, "Application of logical regression function model in credit business of commercial banks," *Appl. Math. Nonlinear Sci.*, vol. 7, no. 1, pp. 513–522, Jan. 2022.
[38] S. Guo, W. Peng, and H. Li, "Research on the default risk and credit strategy of small and medium-sized enterprise," in *Proc. Adv. Econ., Bus. Manage. Res.*, 2021, pp. 281–285.
[39] N. Nehrebecka, "Predicting the default risk of companies. Comparison of credit scoring models: LOGIT vs support vector machines," *Econometrics*, vol. 22, no. 2, pp. 54–73, 2018.
[40] D. Huang, "Evaluation analysis of credit risk for listed real estate companies based on logistic model," in *Proc. 3rd Int. Conf. Educ., Manage. Sci. Econ. (ICEMSE)*, 2019, pp. 518–522.
[41] F. Abdullayeva and S. Suleymanzade, "Cyber security attack recognition on cloud computing networks based on graph convolutional neural network and graphsage models," *Results Control Optim.*, vol. 15, Jun. 2024, Art. no. 100423, doi: 10.1016/j.rico.2024.100423.
[42] F. Abdullayeva and S. Suleymanzade, "Using transfer adaptation method for dynamic features expansion in multi-label deep neural network for recommender systems," *Statist., Optim. Inf. Comput.*, vol. 12, no. 2, pp. 524–529, Feb. 2024, doi: 10.19139/soic-2310-5070-1836.
[43] F. C. Abdullayeva, "Development of collaborative risk assessment method for dynamic federation of clouds," *Problems Inf. Technol.*, pp. 46–58, Jul. 2014.
[44] E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *J. Finance*, vol. 23, no. 4, p. 589, Sep. 1968.
[45] W. Qiu, S. Rudkin, and P. Dłotko, "Refining understanding of corporate failure through a topological data analysis mapping of Altman's Z-score model," *Expert Syst. Appl.*, vol. 156, Oct. 2020, Art. no. 113475.
[46] S. Khemakhem, F. Ben Said, and Y. Boujelbene, "Credit risk assessment for unbalanced datasets based on data mining, artificial neural network and support vector machines," *J. Model. Manage.*, vol. 13, no. 4, pp. 932–951, Nov. 2018.
[47] P. Dłotko, "Ball mapper: A shape summary for topological data analysis," 2019, *arXiv:1901.07410*.
[48] Y. Zhu, C. Xie, G.-J. Wang, and X.-G. Yan, "Comparison of individual, ensemble and integrated ensemble machine learning methods to predict China's SME credit risk in supply chain finance," *Neural Comput. Appl.*, vol. 28, no. S1, pp. 41–50, Dec. 2017.
[49] A. Keramatfar, M. Rafiee, and H. Amirkhani, "Graph neural networks: A bibliometrics overview," *Mach. Learn. With Appl.*, vol. 10, Dec. 2022, Art. no. 100401, doi: 10.1016/j.mlwa.2022.100401.
[50] G. Gogoshin and A. S. Rodin, "Graph neural networks in cancer and oncology research: Emerging and future trends," *Cancers*, vol. 15, no. 24, p. 5858, Dec. 2023, doi: 10.3390/cancers15245858.
[51] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.
[52] R. Offringa, L. Tsai, and S. Witte, "Personal and financial risk typologies among women who engage in sex work in mongolia: A latent class analysis," *Arch. Sexual Behav.*, vol. 46, Aug. 2017, doi: 10.1007/s10508-016-0824-1.
[53] C. Pozzi, F. Ferrario, B. Visciano, and L. D. Vecchio, "Corticosteroids in patients with IGA nephropathy and severe chronic renal damage," *Case Rep. Nephrology*, vol. 2012, no. 1, 2012, Art. no. 180691.
[54] T. Millington and M. Niranjan, "Construction of minimum spanning trees from financial returns using rank correlation," *Physica A, Stat. Mech. Appl.*, vol. 566, Mar. 2021, Art. no. 125605, doi: 10.1016/j.physa.2020.125605.

KOSAR FARAJPOUR MOJDEHI received the B.Sc. degree in mathematics from Alzahra University, Tehran, Iran, in 2020, and the M.Sc. degree in industrial engineering with a focus on information systems from Iran University of Science and Technology, Tehran, in 2024. Her research interests include credit risk assessment, machine learning algorithms, and topological data analysis.

BABAK AMIRI received the Ph.D. degree in information technology from The University of Sydney, in 2014. He is currently an Assistant Professor with Iran University of Science and Technology. He has a track record of publishing quality research articles in respected outlets within the fields of complex systems, machine learning, and data science. Notably, the academic community has well-received his research on topics, such as social media analysis, complex networks, customer analytics, and text analysis.

AMIRALI HADDADI received the degree in industrial engineering from Iran University of Science and Technology, Tehran, Iran. He has experience working in the machine learning and AI industry, contributing to various projects and companies. His research interests include machine learning, text mining, and financial markets.