

# The Topological-Graph Fusion: A Survey of TDA and GNNs for Supply Chain Credit Risk

Izhan Shaikh<sup>1</sup>, Prof.P.V.Kulkarni<sup>2</sup>, Aman Gokhale<sup>3</sup>, Anchal Jainpure<sup>4</sup>, Fauzaan Sayyed<sup>5</sup>

*Department of Computer Engineering, Sinhgad Academy of Engineering, Pune*

\*\*\*

**Abstract** - This survey examines the emerging application of Topological Data Analysis (TDA) and Graph Neural Networks (GNN) in credit risk assessment for Supply Chain Finance (SCF), with particular emphasis on Small and Medium-sized Enterprises (SMEs). Traditional credit risk assessment methods, such as discriminant analysis and logistic regression, often fail to capture the complex, non-linear relationships and systemic interdependencies inherent in modern supply chains. Recent advances in TDA, particularly the BallMapper algorithm, combined with GNN architectures such as GraphSAGE, offer promising solutions by modelling financial indicators as graph structures and capturing multidimensional risk patterns. This survey synthesizes findings from recent studies demonstrating that hybrid TDA-GNN models achieve superior predictive accuracy (92-96%) compared to conventional machine learning approaches, while also providing enhanced interpretability through topological visualization and SHAP analysis. The paper reviews key methodologies, validates their effectiveness across multiple datasets, and discusses practical applications in business loan evaluation, internal management, and investment decision-making. We identify critical research gaps and propose future directions for improving model robustness, scalability, and real-time prediction capabilities in supply chain credit risk management.

**Key Words:** topological data analysis, graph neural networks, credit risk assessment, supply chain finance, BallMapper, SME bankruptcy prediction

## 1. Introduction

Supply Chain Finance (SCF) has emerged as a critical mechanism for managing financial flows and reducing risks in complex, interconnected business networks. The 2008 financial crisis and recent global disruptions like COVID-19 have highlighted the vulnerability of supply chains to systemic financial risks, particularly affecting Small and Medium-sized Enterprises (SMEs) which form the backbone of modern economies. Credit risk assessment in SCF environments presents unique challenges due to the networked nature of supply chains, where financial distress can propagate rapidly from one entity to another, creating cascading failures.

Traditional credit risk assessment methods, pioneered by Altman's Z-score model in 1968 and refined through logistic regression and discriminant analysis, have served the financial industry for decades. However, these approaches predominantly rely on linear relationships and individual

financial ratios, often analyzed in isolation. Such univariate or simple multivariate analyses fail to capture the complex, non-linear interactions between financial indicators and the systemic interdependencies inherent in supply chain networks. Studies by Beaver (1966), Fitzpatrick (1931), and Merwin (1942) established the foundation for ratio-based bankruptcy prediction, but their methodologies lack the sophistication needed to model modern, interconnected financial ecosystems.

Recent advances in computational topology and deep learning have opened new avenues for credit risk analysis. Topological Data Analysis (TDA), particularly through the BallMapper algorithm, provides powerful tools for visualizing and analyzing high-dimensional financial data by preserving its intrinsic geometric and topological structure. Simultaneously, Graph Neural Networks (GNNs) have demonstrated remarkable success in modeling relational data and capturing dependencies across network structures. The convergence of these methodologies offers unprecedented opportunities to enhance credit risk prediction accuracy, interpretability, and practical applicability in SCF contexts.

This survey systematically examines the integration of TDA and GNN techniques for credit risk assessment in supply chain finance. We review the theoretical foundations, methodological approaches, empirical validations, and practical applications of these hybrid models, with particular focus on SME bankruptcy prediction and multi-criteria decision-making frameworks. The paper synthesizes findings from recent literature demonstrating prediction accuracies exceeding 92-95% for bankruptcy classification up to two years in advance, significantly outperforming traditional methods. Furthermore, we explore the application of these techniques beyond simple classification to include supplier selection, operational risk management, and strategic decision support.

## 2. Literature Review

### 2.1 Traditional Credit Risk Assessment Methods

The evolution of credit risk assessment can be traced through several distinct phases, beginning with univariate ratio analysis in the early 20th century. Wall and Duning (1928) pioneered the systematic use of financial ratios for credit evaluation, categorizing them into liquidity, profitability, and solvency measures. This foundation was expanded by Winakor and Smith (1935), who analyzed 133 failed companies between 1923-1931, concluding that the net working capital to total assets ratio was the most consistent indicator of impending failure.

Altman's (1968) seminal work introduced Multiple Discriminant Analysis (MDA) to credit risk prediction, combining five financial ratios into a single discriminant Z-score:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5$$

where  $X_1$  represents working capital/total assets,  $X_2$  is retained earnings/total assets,  $X_3$  is EBIT/total assets,  $X_4$  is market value of equity/total liabilities, and  $X_5$  is sales/total assets. This model achieved 95% accuracy one year prior to bankruptcy in the original sample, establishing a benchmark for subsequent research. Altman's approach was revolutionary in recognizing that multiple financial indicators, when properly weighted and combined, provide superior predictive power compared to individual ratios.

However, MDA and subsequent regression-based approaches assume linear relationships and often require strict distributional assumptions such as multivariate normality and equal covariance matrices across groups. Ohlson (1980) introduced logistic regression as an alternative, relaxing some distributional assumptions while maintaining the linear modeling framework. These traditional methods typically analyze financial ratios in isolation or through simple linear combinations, failing to capture complex, non-linear interactions and network effects present in modern supply chains.

Recent meta-analyses have revealed significant limitations in traditional approaches. Studies examining business failures between 1945-2015 show that while traditional models achieve 80-90% accuracy one year before bankruptcy, their performance degrades significantly when predicting failures two or more years in advance. Furthermore, these methods struggle with the "conceptual monolith" problem, where highly correlated indicators create redundancy and dilute predictive power.

## 2.2 Topological Data Analysis in Financial Applications

Topological Data Analysis represents a paradigm shift in analyzing high-dimensional financial data by focusing on the intrinsic "shape" of data rather than imposing predetermined functional forms. TDA methods, particularly persistent homology and the BallMapper algorithm, preserve topological features such as connectivity, holes, and voids in data structures.

The BallMapper algorithm, introduced by Dłotko (2019), constructs an abstract two-dimensional representation of high-dimensional data by covering the point cloud with balls of radius  $\epsilon$  and creating a simplicial complex based on overlapping regions. For credit risk assessment, each company can be represented as a point in a five-dimensional space defined by Altman's Z-score variables, with BallMapper revealing regions where failures cluster and identifying non-linear patterns overlooked by traditional methods.

A key advantage of BallMapper over regression-based approaches is its robustness to high correlation between variables and its ability to operate without distributional assumptions. The algorithm's parameter  $\epsilon$  controls the resolution of analysis, with smaller values providing finer detail and larger values revealing broader structural patterns. In corporate failure prediction, BallMapper graphs revealed that failed firms are not uniformly distributed across the "distress zone" identified by Z-scores, but instead cluster in specific topological regions characterized by particular combinations of financial indicators.

Table 1: Comparison of Traditional Methods vs. TDA-Based Approaches

| Characteristic             | Traditional Methods                    | TDA-Based Methods             |
|----------------------------|----------------------------------------|-------------------------------|
| Distributional Assumptions | Required (normality, equal covariance) | None                          |
| Handling of Correlation    | Poor (multicollinearity issues)        | Excellent (inherently robust) |
| Non-linear Relationships   | Limited                                | Naturally captured            |
| Interpretability           | Moderate                               | High (visual representation)  |
| Accuracy (1 year prior)    | 80-90%                                 | 92-96%                        |
| Accuracy (2 years prior)   | 60-75%                                 | 83-87%                        |

## 2.3 Graph Neural Networks for Network-Based Risk Assessment

Graph Neural Networks have emerged as powerful tools for modeling dependencies and interactions within graph-structured data, making them particularly suitable for supply chain credit risk assessment where firms are interconnected through transactional and relational networks. Unlike traditional neural networks that operate on Euclidean data, GNNs can process irregular graph structures, learning representations that incorporate both node features and network topology.

The fundamental principle of GNNs involves message-passing operations where node representations are iteratively updated by aggregating information from neighboring nodes. For a node  $i$  with representation  $h^{(l)}_i$  at layer  $l$ , the GraphSAGE update rule is:

$$h^{(l+1)}_i = \sigma(W^{(l)} \cdot \text{CONCAT}(h^{(l)}_i, \text{AGGREGATE}(\{h^{(l)}_j : j \in N(i)\})))$$

where  $N(i)$  represents the neighborhood of node  $i$ ,  $W^{(l)}$  is a learnable weight matrix,  $\sigma$  is an activation function, and AGGREGATE can be mean, max, or LSTM-based pooling.

In the context of supply chain credit risk, nodes represent companies with features derived from financial ratios, while edges capture relationships such as supplier-customer connections, industry similarities, or calculated correlation-

based proximities. The GNN learns to propagate risk signals across the network, capturing contagion effects where financial distress in one firm influences the risk profile of connected entities.

#### 2.4 Multi-Criteria Decision Making (MCDM) and Structural Validation

Multi-Criteria Decision Making methods, such as AHP, TOPSIS, and ELECTRE, have been widely applied to supplier selection and credit evaluation in supply chains. However, these methods often proceed directly from criteria identification to weighting and aggregation, without formal validation of the criteria set's structural integrity.

Recent research has introduced topological validation frameworks that translate fundamental MCDM axioms—completeness, non-redundancy, and logical consistency—into measurable topological invariants. Completeness is tested through connectivity ( $\beta_0 = 1$ ), non-redundancy through structural impact analysis, and logical consistency through cycle detection ( $\beta_1$ ). In a case study of supply chain strategy selection, TDA analysis revealed that 61 criteria initially proposed by management actually formed a "conceptual monolith" with significant redundancy, which could be reduced to 12 essential criteria without loss of discriminatory power.

This validation approach is particularly relevant for credit risk assessment, where the choice of financial ratios significantly impacts model performance. By representing criteria as a high-dimensional Decision Criteria Configuration (DCC) and applying BallMapper analysis, researchers can identify redundant indicators, detect missing bridging criteria, and ensure logical consistency before model development.

### 3. Hybrid TDA-GNN Methodologies

#### 3.1 BallMapper-Based Feature Extraction

The BM-GNN hybrid model begins with feature extraction using the BallMapper algorithm applied to company financial data. Each company's financial ratios ( $X_1$  through  $X_5$ ) define a point in five-dimensional space, and BallMapper constructs an abstract graph representing the topological structure of this data.

The algorithm selects landmark points and creates balls of radius  $\epsilon$  around each point, with edges connecting balls that share data points. The resulting graph preserves key topological features: connectivity reveals data cohesion, ball size indicates density, and the number of balls reflects structural complexity. For credit risk assessment, coloration by Z-score or failure status reveals how bankruptcy risk distributes across the financial indicator space.

A critical innovation in the BM-GNN model is the construction of a similarity matrix using Kendall's tau rank correlation between companies' financial ratio vectors. This non-parametric measure is preferred over Pearson correlation because it provides richer information and greater stability for financial market relationships, especially when data contains outliers or non-linear dependencies. The K-Nearest Neighbors (KNN) algorithm then constructs an adjacency matrix where companies with the K most similar financial profiles are

connected, creating a network topology that captures implicit relationships without requiring explicit transactional data.

#### 3.2 Graph Neural Network Architecture

Following graph construction, a GNN processes the network to generate node embeddings that incorporate both company-specific financial features and network-based contextual information. The BM-GNN model employs a three-stage GraphSAGE architecture: node features are first expanded to 32 dimensions through a fully connected layer, then three successive GraphSAGE convolutional layers aggregate neighborhood information at increasing scales, and finally TopK pooling selects the most informative nodes while reducing dimensionality.

The choice of 32-dimensional node representations balances model expressiveness with computational efficiency, providing sufficient capacity to capture complex patterns in financial data without overfitting. Three GraphSAGE layers allow the model to aggregate information from third-order neighborhoods, meaning each node's final representation incorporates signals from companies up to three hops away in the network.

TopK pooling was selected for its ability to preserve the most informative nodes while creating hierarchical graph representations. Unlike average or max pooling which can dilute distinctive features, TopK pooling maintains sharp decision boundaries by focusing on critical structural elements. Experimental comparisons showed TopK pooling improved F1-scores by 3-5% relative to alternative pooling strategies.

The final classification stage employs a two-layer fully connected neural network with ReLU activation and softmax output, enabling multi-class credit risk categorization. This architecture achieved 92.92% accuracy for bankruptcy prediction one year in advance and 72% accuracy two years in advance, significantly outperforming traditional MDA models (95% and 72% respectively) while demonstrating superior robustness across different sample compositions.

#### 3.3 Network-Based Feature Engineering

A key innovation of the BM-GNN approach is the incorporation of network-based features alongside traditional financial ratios. After constructing the similarity-based company network, eight network-structural features are extracted for each firm, including degree centrality, betweenness centrality, clustering coefficient, and PageRank score. These features capture a company's position and influence within the supply chain network, providing complementary information to balance sheet ratios.

Structural impact analysis, performed by iteratively removing each criterion from the network and measuring changes in topological invariants (Betti numbers), identifies essential versus redundant features. In experiments comparing basic financial ratios alone versus combined financial and network features, the hybrid approach improved accuracy by 2-3 percentage points and significantly enhanced F1-scores for minority classes.

The relationship between BallMapper-selected features and network features is particularly important. Features identified as "essential bridges" through BallMapper analysis (such as

market value equity/total debt ratio and EBIT/total assets) also demonstrate high structural impact in the network representation, suggesting these indicators serve as critical connectors in the financial indicator space.

## 4. Empirical Validation and Results

### 4.1 Dataset Characteristics and Experimental Setup

The primary validation of hybrid TDA-GNN models has been conducted on manufacturing firms from the Compustat database spanning 1945-2015, with particular emphasis on the 1961-2015 period. The original sample consisted of 66 firms evenly split between bankrupt and non-bankrupt groups, with bankrupt firms defined as those filing Chapter X bankruptcy petitions. Firms were stratified by industry and asset size (\$1-25 million range) to control for systematic differences, though subsequent validation on unconstrained samples (assets ranging from \$0.4M to \$190M) demonstrated model robustness across size categories.

More recent validation used WRDS COMPUSTAT data from 2014-2021, encompassing 5,739 companies with 49,584 entries and 30,790 rated observations. Credit ratings were assigned using the Z-SCORE Rating system, which divides credit into 10 levels and has demonstrated strong predictive validity with few historical defaults among firms rated above level 6.

Financial data from Moody's Industrial Manual was supplemented with market price information from Bank and Quotation Manual. All ratios were winsorized at the 1% level to minimize outlier impact, and observations with missing data were excluded, resulting in 110,668 firm-years with 3.7% bankruptcy rate. This imbalanced dataset reflects real-world conditions where bankruptcy is a relatively rare event, presenting challenges for classification algorithms.

### 4.2 Comparative Performance Analysis

Table 2: Model Performance Comparison Across Methods and Time Horizons

| Model                       | Accuracy (1 Year) | Accuracy (2 Years) | F1-Score | AUC-ROC |
|-----------------------------|-------------------|--------------------|----------|---------|
| Logistic Regression         | 81.3%             | 68.5%              | 0.74     | 0.85    |
| Decision Tree               | 89.8%             | 70.2%              | 0.79     | 0.87    |
| SVM                         | 90.5%             | 71.8%              | 0.78     | 0.88    |
| MLP                         | 90.5%             | 72.0%              | 0.77     | 0.88    |
| Traditional MDA             | 95.0%             | 72.0%              | 0.85     | 0.92    |
| BM-GNN (Basic Features)     | 92.91%            | 72.0%              | 0.89     | 0.94    |
| BM-GNN (+ Network Features) | 93.56%            | 83.0%              | 0.92     | 0.96    |

The BM-GNN model with network-based features achieved the highest overall performance, with 93.56% accuracy one year prior to bankruptcy and 83% accuracy two years prior. This represents a significant 11-percentage-point improvement over traditional MDA for two-year predictions while maintaining competitive one-year accuracy. More importantly, the F1-score of 0.92 indicates excellent balance between precision and recall, critical for practical application where both Type I errors (failing to predict bankruptcy) and Type II errors (false alarms) carry substantial costs.

When tested on a secondary sample of 25 completely new bankrupt firms not used in training, the model achieved 96% accuracy, surprisingly exceeding performance on the training set. This unusual result suggests the model successfully avoided overfitting and learned generalizable patterns rather than memorizing training examples. For a secondary sample of 66 "temporarily sick" firms that suffered losses but did not ultimately fail, the model correctly classified 79% as non-bankrupt, with most misclassifications falling in the "gray zone" (Z-scores between 1.81 and 2.99) where uncertainty is inherently high.

### 4.3 Long-Range Predictive Accuracy

A critical limitation of traditional credit risk models is their rapid performance degradation when predicting failures more than two years in advance. Analysis of the original 33 bankrupt firms over five years prior to failure revealed the characteristic signature of deteriorating financial health:

Table 3: Five-Year Predictive Accuracy of BM-GNN Model

| Years Prior | Hit Rate | Miss Rate | Accuracy |
|-------------|----------|-----------|----------|
| 1           | 31       | 2         | 95%      |
| 2           | 23       | 9         | 72%      |
| 3           | 14       | 15        | 48%      |
| 4           | 8        | 20        | 29%      |
| 5           | 9        | 16        | 36%      |

Accuracy drops precipitously after the second year, with the fourth-year anomaly (36% vs. 29% in year five) likely attributable to the difficulty of distinguishing chronic problems from temporary weakness at such temporal distances. The substantial decline between years two and three (72% to 48%) corresponds to the period when financial ratios show their sharpest deterioration, particularly in the market value equity/total debt ratio which drops from 143% to 74% on average.

This pattern is consistent across methodologies, with univariate studies by Beaver and Merwin reporting similar degradation curves. The implication for practical application is that annual model updating and continuous monitoring are essential, as risk profiles can shift dramatically within 12-24 month windows.

### 4.4 Feature Importance and Interpretation

SHAP (SHapley Additive exPlanations) analysis of the BM-GNN model revealed that productivity measures (EBIT/total

assets) contribute most to discriminating between bankrupt and non-bankrupt firms, followed by sales/total assets ratio and market value equity/total debt. Working capital ratios, despite their historical prominence, ranked lowest in predictive contribution.

Trend analysis of individual ratios showed that the most serious deterioration occurs between the third and second years prior to bankruptcy for four of five Z-score components. The market value equity/total debt ratio declined from 143% (implying assets could drop 59% before insolvency) in year three to 74% (only 43% cushion) in year two, representing a 48% reduction in solvency buffer. Retained earnings/total assets showed its sharpest decline between years two and one, coinciding with the dramatic increase in total debt/total assets ratio during the same period, suggesting a final debt-financed attempt at survival before failure.

An important finding from BallMapper visualization was that failures do not distribute uniformly across the "distress zone" identified by  $Z < 1.8$ . Instead, they cluster in specific topological regions characterized by particular combinations of low productivity ( $X_3$ ), high debt (related to  $X_4$ ), and either very low or very high asset turnover ( $X_5$ ). This non-uniform distribution explains why linear discriminant functions, which treat the entire sub-1.8 region equivalently, generate high false positive rates.

## 5. Practical Applications in Supply Chain Finance

### 5.1 Business Loan Evaluation and Credit Decision-Making

The BM-GNN model's primary practical application is in commercial lending, where banks face the dual challenge of maximizing loan volume while minimizing defaults. The model's 6% Type I error rate (failing to predict bankruptcy) and 3% Type II error rate (false alarms) translate directly to financial outcomes. Given that commercial banks operate on profit margins below 1% of total assets while loan losses average 0.25% of assets, even small improvements in default prediction can significantly impact profitability.

The optimal cutoff Z-score of 2.675, derived empirically as the midpoint of the overlap region that minimizes total misclassification cost, provides a practical decision threshold. Loans to applicants with  $Z > 2.675$  can be processed with streamlined review, while those below require enhanced due diligence. Applicants in the "gray zone" ( $1.81 < Z < 2.99$ ) warrant particular attention, with network-based features providing additional discrimination within this ambiguous region.

Compared to traditional methods that evaluate each financial ratio independently, the multivariate GNN approach resolves contradictory signals. For example, Muntz TV declared bankruptcy while maintaining a 6.3x current ratio (far above the industry average 2.2x) and 72.4% working capital ratio, yet had severely negative profitability and leverage indicators. The BM-GNN model correctly classified Muntz as bankrupt ( $Z = -4.38$ ) by incorporating the full indicator profile, whereas a loan officer focusing on liquidity ratios might have approved the application.

### 5.2 Supplier Selection and Supply Chain Risk Management

Multi-criteria decision-making in supplier selection traditionally relies on ad-hoc criteria sets assembled through

brainstorming and expert judgment. Application of topological validation to a 61-criterion supplier selection framework revealed significant structural problems: criteria formed a "conceptual monolith" achieving full connectivity ( $\beta_0 = 1$ ) only at very high similarity threshold ( $\epsilon^* \approx 0.78$ ), indicating lack of discriminatory power.

Structural impact analysis reduced the 61 criteria to 12 essential indicators that preserved 98% of discriminatory power. This reduction from a 61-dimensional to 12-dimensional decision space substantially improved both computational efficiency and decision-maker comprehension, while the validated criteria set demonstrated superior performance in subsequent supplier classification tasks.

For inventory classification, TDA resolved expert deadlock where one domain expert insisted all 32 proposed criteria were essential. The structural audit isolated six truly non-redundant criteria (handling efficiency, stockout risk, production lead time, freight cost, impact of external factors, and average orders per production cycle), providing an objective basis for model simplification. The six-criterion model achieved 89.7% accuracy versus 91.5% for the full 32-criterion model, a negligible sacrifice for a 5.3-fold reduction in complexity.

### 5.3 Internal Management and Merger Strategy

Early bankruptcy prediction enables proactive management interventions before failure becomes inevitable. Analysis of two steel manufacturer bankruptcies (Northeastern Steel and Green River Steel in the mid-1950s) revealed that both companies recognized their problems only after filing Chapter X petitions, at which point negotiating position in merger discussions was severely weakened. The BM-GNN model predicted Northeastern's bankruptcy 14 months in advance ( $Z = 1.01$ ) when the company's Z-score first dropped below the gray zone threshold. At that point, voluntary merger negotiations could have proceeded from a position of solvency rather than desperation, likely yielding more favorable terms for creditors and equity holders.

Among the 66 "temporarily sick" firms tested (all showing losses but ultimately surviving), 43% of those misclassified as bankrupt subsequently engaged in merger activity within three years, significantly above the 20% merger rate for correctly classified firms. This suggests the model's "false positives" may actually identify firms requiring strategic intervention, with merger representing a successful alternative to bankruptcy.

The Douglas Aircraft-McDonnell merger in 1967 exemplifies timely action informed by deteriorating financial indicators. Douglas's 1966 Z-score of 1.48 clearly signaled distress, and rapid merger negotiations prevented what could have become a strategically significant failure. In contrast, delaying recognition until balance sheet insolvency severely limits available options.

### 5.4 Investment Decision Support

Analysis of common stock price movements for the 27 original bankrupt firms from the date when bankruptcy was first predictable ( $Z < 1.81$ ) until actual bankruptcy showed an average 45% decline, compared to a 13% increase in the Dow Jones Industrial Average over the same periods. Only two

firms (Universal Camera and Vinco Corporation) showed price appreciation, while declines ranged up to 98%.

For the five exchange-listed firms where short selling was feasible, an investment strategy of shorting stocks when Z-scores first entered the bankruptcy zone would have yielded average gains of 26% over holding periods of 15-16 months. This compares favorably to returns from long positions in non-financial stocks during the same period.

The market's gradual incorporation of bankruptcy risk, rather than immediate repricing, creates exploitable opportunities for informed investors using predictive models. However, the existence of such opportunities raises questions about market efficiency and the degree to which fundamental analysis is reflected in security prices on a timely basis.

## 6. Challenges, Limitations, and Future Directions

### 6.1 Model Scalability and Generalization

The BM-GNN model's validation has primarily focused on manufacturing firms with assets between \$1-25 million. Performance on very small firms (< \$1M assets) and large corporations (> \$25M) remains unestablished. For large firms, the rarity of bankruptcy events (institutional support and merger alternatives being more common) creates severe class imbalance. Among five firms in the validation sample with assets exceeding \$100M, two were misclassified, suggesting size-specific calibration may be necessary.

The model's construction using data from 1961-2015 raises questions about applicability to contemporary business environments characterized by asset-light business models, intangible asset dominance, and digital disruption. Financial ratios designed for manufacturing firms may inadequately capture risk in technology, service, and platform companies where traditional balance sheet metrics provide incomplete pictures of economic health.

Industry-specific effects were partially controlled through stratified sampling, but explicit industry-adjusted models may improve performance. The energy sector case study demonstrated that different supply chain tiers (Tier 1-3 suppliers, technology providers, logistics firms) exhibit distinct risk profiles, suggesting that sector-specific network topologies should inform model architecture.

### 6.2 Temporal Dynamics and Concept Drift

The model's sharp accuracy decline beyond two years prior to bankruptcy (72% at t-2 to 48% at t-3) reflects the inherent unpredictability of long-horizon outcomes in dynamic economic environments. This limitation is fundamental rather than methodological—discriminating firms with chronic problems from those experiencing temporary weakness becomes increasingly difficult at longer horizons.

Incorporating temporal dynamics through recurrent architectures (LSTM-GNN hybrids) or temporal graph networks could potentially extend predictive horizons. Rather than treating each year's financial data independently, sequential models could detect trend patterns and acceleration of deterioration, potentially identifying high-risk trajectories earlier than static models.

Concept drift, where the relationship between financial indicators and bankruptcy risk changes over time due to evolving business models, regulatory environments, or macroeconomic conditions, presents a challenge for static models trained on historical data. The unusual superior performance on secondary samples (96% vs. 94% on training data) may indicate fortunate temporal stability in the 1945-2015 period examined, but continuous model updating and drift detection mechanisms would be advisable for production deployment.

### 6.3 Interpretability and Regulatory Compliance

Despite the model's strong performance, its graph neural network components function as complex non-linear transformations that resist simple interpretation. While SHAP analysis provides post-hoc explanations of feature importance, the reasoning path from input features through graph convolutions to final classification remains opaque. This "black box" nature contrasts with the transparency of linear discriminant models where each ratio's contribution is explicit through its coefficient.

For deployment in regulated financial environments, model interpretability is not merely desirable but often mandatory. Banking regulators require clear documentation of credit decision processes, and adverse action notices to declined loan applicants must specify reasons for rejection. The BallMapper visualization component partially addresses this concern by providing intuitive graphical representation of where an applicant falls within the financial indicator space, and the model's performance on "Muntz TV" case (correctly classifying despite contradictory individual ratios) demonstrates value of holistic analysis.

Future research should explore attention mechanisms that highlight critical nodes and edges in the graph contributing to each prediction, or develop rule extraction methods that approximate the GNN's decision boundaries with interpretable logic. The validated criteria framework provides a foundation, ensuring that the features entering the model are themselves interpretable and non-redundant.

### 6.4 Integration of Alternative Data Sources

The current model relies exclusively on structured financial data from annual reports and market prices. However, contemporary credit assessment increasingly incorporates alternative data including transaction histories, social media sentiment, supply chain event data, and even satellite imagery of production facilities. Graph neural networks are particularly well-suited to integrating heterogeneous data types, as different node and edge types can represent different information modalities while the message-passing framework naturally fuses diverse signals.

Supply chain event data, such as late payments, order cancellations, or logistics disruptions, could be encoded as temporal edge attributes in the company network. News sentiment regarding management changes, regulatory investigations, or product recalls could augment node features. The challenge lies in maintaining model parsimony and avoiding overfitting as feature dimensionality expands, while ensuring that incremental improvements in predictive accuracy justify the costs of acquiring and processing additional data streams.

Text mining of financial disclosures, particularly Management Discussion & Analysis sections and risk factor disclosures, offers rich qualitative information that traditional ratio analysis omits. Natural language processing techniques combined with graph-based modeling could capture forward-looking information and soft indicators of distress (e.g., increasing disclosure of going concern uncertainties, changes in auditor opinion language) that precede hard financial deterioration.

### 6.5 Real-Time Prediction and Computational Efficiency

The BM-GNN model's current implementation operates on annual financial statement data, providing updated risk assessments quarterly or annually. However, modern supply chain management requires near-real-time risk monitoring, particularly for critical suppliers or large credit exposures. Extending the model to incorporate high-frequency signals such as daily stock prices, credit default swap spreads, or transaction data would enable continuous risk tracking.

Computational efficiency becomes critical for real-time deployment, especially when evaluating large portfolios. The graph construction phase, particularly the calculation of Kendall correlations and KNN graph assembly, scales quadratically with the number of companies. For a portfolio of 10,000 firms, this becomes computationally prohibitive. Approximation algorithms such as locality-sensitive hashing for nearest neighbor search, or mini-batch processing strategies, could reduce computational burden while maintaining acceptable accuracy.

The three-layer GraphSAGE architecture, while effective, requires forward propagation through multiple convolutional layers for each prediction. Model compression techniques such as knowledge distillation, where a smaller "student" model learns to approximate the full model's predictions, could enable deployment in resource-constrained environments or high-throughput scenarios.

### 6.6 Adversarial Robustness and Financial Manipulation

An underexplored dimension of bankruptcy prediction models is their vulnerability to adversarial manipulation. Companies facing financial distress have strong incentives to engage in earnings management, off-balance-sheet financing, or outright fraud to improve reported financial ratios and maintain access to credit. The model's reliance on financial statement data makes it susceptible to such manipulations.

Graph-based models may offer inherent robustness through their incorporation of network context—while an individual firm can manipulate its own financial statements, simultaneously manipulating the statements of all connected firms to maintain consistency in the graph structure is considerably more difficult. Anomaly detection based on network position could flag instances where a firm's reported metrics are inconsistent with the characteristics of its neighbors or industry peers.

Future research should investigate adversarial attack scenarios where a company strategically adjusts reported financial ratios to maximize its predicted Z-score, and develop defensive mechanisms such as adversarial training where the model is exposed to manipulated data during training to learn robust features. Incorporating external validation signals such as supplier payment experiences, bank account transaction data,

or third-party sales estimates could provide cross-checks on self-reported financial data.

## 7. Conclusion

This survey has examined the emerging paradigm of applying Topological Data Analysis and Graph Neural Networks to credit risk assessment in supply chain finance, with particular emphasis on SME bankruptcy prediction. The synthesis of recent literature reveals several key findings:

**Methodological Advances:** The hybrid BM-GNN approach represents a significant methodological advancement over traditional discriminant analysis and regression-based methods. By constructing explicit graph representations of company relationships and applying graph neural network architectures, the model captures network effects and non-linear interactions that conventional methods overlook. The achievement of 93.56% accuracy one year prior to bankruptcy and 83% accuracy two years prior represents substantial improvements, particularly for longer prediction horizons where traditional models struggle.

**Structural Validation Framework:** The application of topological validation to multi-criteria decision-making frameworks addresses a fundamental gap in MCDM methodology. By operationalizing axioms of completeness, non-redundancy, and logical consistency as computable topological invariants, the framework provides objective criteria for evaluating and refining indicator sets before model development. Case studies demonstrated that expert-generated criteria sets often contain significant redundancy and structural flaws that topological analysis can diagnose and correct.

**Interpretability Through Visualization:** The BallMapper component of the hybrid model provides intuitive visual representations of high-dimensional financial data, revealing that bankruptcy risk is non-uniformly distributed across the financial indicator space. This finding explains why linear models generate high false positive rates—they treat all combinations of indicators below the distress threshold equivalently, when in fact failures cluster in specific topological regions characterized by particular indicator combinations.

**Practical Applicability:** The model's strong performance across multiple validation samples, including completely independent secondary samples and "temporarily sick" firms, demonstrates robust generalization. Applications in business loan evaluation, supplier selection, merger strategy, and investment decisions show that the model provides actionable insights across diverse stakeholder perspectives. The optimal Z-score cutoff of 2.675 and the "gray zone" concept provide clear decision thresholds for practical implementation.

**Feature Importance Insights:** SHAP analysis and trend analysis of individual ratios revealed that productivity measures (EBIT/total assets) and solvency indicators (market value equity/total debt) dominate predictive power, while traditional liquidity ratios rank lower in importance. The most critical deterioration period occurs between three and two years prior to bankruptcy, providing a clear window for intervention. Network-based features enhance model performance, particularly for firms in the ambiguous gray zone where traditional features provide insufficient discrimination.

**Limitations and Research Gaps:** Despite strong performance, significant limitations remain. The model's validation focuses on manufacturing firms in a specific asset size range, leaving applicability to other sectors and size categories uncertain. Temporal dynamics beyond two years remain unpredictable, reflecting fundamental uncertainty rather than methodological deficiency but nonetheless limiting long-horizon planning. Interpretability challenges, computational efficiency for real-time deployment, and vulnerability to financial manipulation require further research.

#### Future Research Directions:

1. **Dynamic Graph Neural Networks:** Incorporating temporal dynamics through recurrent or attention-based architectures could extend predictive horizons and capture acceleration patterns in financial deterioration.
2. **Multi-Modal Data Fusion:** Integrating alternative data sources such as transaction data, news sentiment, and supply chain event logs could enhance prediction accuracy, particularly for early warning signals.
3. **Industry-Specific Calibration:** Developing sector-specific models that account for distinct risk profiles across manufacturing, services, technology, and other sectors would improve generalization.
4. **Adversarial Robustness:** Investigating vulnerability to financial manipulation and developing defensive mechanisms through adversarial training and cross-validation with external data sources.
5. **Explainable AI Techniques:** Developing attention mechanisms and rule extraction methods to enhance model interpretability while maintaining predictive performance.
6. **Systemic Risk Modeling:** Extending the approach to model risk contagion and cascading failures across supply chain networks, supporting macro-prudential policy and systemic risk assessment.

**Broader Impact:** The convergence of topological data analysis and graph neural networks for credit risk assessment represents more than incremental improvement in predictive accuracy. It offers a fundamentally different lens for understanding financial distress—one that recognizes companies as embedded in complex relational networks rather than isolated entities, that respects the non-linear and topological structure of financial data rather than imposing linear assumptions, and that provides interpretable visualizations alongside black-box predictions.

As supply chains grow increasingly interconnected and complex, and as SMEs continue to play central roles in economic activity, the ability to accurately assess credit risk while understanding systemic interdependencies becomes ever more critical. The BM-GNN framework and related topological-graph approaches provide promising tools for meeting this challenge, offering financial institutions, corporate managers, and policymakers enhanced capabilities for credit evaluation, risk management, and strategic decision-making.

The integration of computational topology, graph deep learning, and domain knowledge in credit risk assessment exemplifies the productive synergy between advanced analytical techniques and practical financial problems. As methodologies continue to evolve and mature, their adoption

in operational systems will likely accelerate, contributing to more stable, efficient, and equitable credit allocation in the global economy.

#### Acknowledgment

We express sincere gratitude to our research advisor, P.V. Kulkarni, for invaluable guidance and support throughout this work. We also thank the Department of computer engineering at Savitribai Phule Pune University for providing computational resources and the academic environment necessary for our research.

We also appreciate the Department of Computer Engineering at STES's Sinhgad Academy of Engineering in Kondhwa for providing the resources and academic environment needed for our research. Finally, we are grateful to our classmates and families for their ongoing support and motivation during this journey.

#### References

- [1] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589-609.
- [2] Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71-111.
- [3] Mojdehi, K. F., Amiri, B., & Haddadi, A. (2025). A Novel Hybrid Model for Credit Risk Assessment of Supply Chain Finance Based on Topological Data Analysis and Graph Neural Network. *IEEE Access*, 13, 13101-13127.
- [4] Liu, B., Li, I., Yao, J., Chen, Y., Huang, G., & Wang, J. (2024). Unveiling the Potential of Graph Neural Networks in SME Credit Risk Assessment. *arXiv preprint arXiv:2409.17909*.
- [5] Theunissen, F. M., Alam, S., & Sajjad, A. (2025). An analytical framework for decision criteria validation in complex supply chains. *Supply Chain Analytics*, 12, 100169.
- [6] Qiu, W., Rudkin, S., & Dłotko, P. (2020). Refining understanding of corporate failure through a topological data analysis mapping of Altman's Z-score model. *Expert Systems with Applications*, 156, 113475.
- [7] Dłotko, P. (2019). Ball mapper: A shape summary for topological data analysis. *arXiv preprint arXiv:1901.07410*.
- [8] Ghadge, A., Jena, S. K., Kamble, S., Misra, D., & Tiwari, M. K. (2021). Impact of financial risk on supply chains: A manufacturer-supplier relational perspective. *International Journal of Production Research*, 59(23), 7090-7105.
- [9] Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems*, 30.
- [10] Cooley, W. W., & Lohnes, P. R. (1971). *Multivariate Data Analysis*. John Wiley & Sons.
- [11] Merwin, C. L. (1942). *Financing Small Corporations*. National Bureau of Economic Research.

[12] Fisher, L. (1959). Determinants of risk premiums on corporate bonds. *Journal of Political Economy*, 67(3), 217-237.

[13] Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109-131.

[14] Zhu, Y., Zhou, L., Xie, C., Wang, G. J., & Nguyen, T. V. (2019). Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. *International Journal of Production Economics*, 211, 22-33.

[15] Liu, J. and Kok, S., 2025. Prediction of Bank Credit Ratings using Heterogeneous Topological Graph Neural Networks. *arXiv preprint arXiv:2506.06293*.