IJOER
ENGINEERING JOURNAL

Engineering Journal IJOER

VOLUME-12, ISSUE-3, MARCH 2026

DOWNLOAD NOW

Contact us +91-7665235235 www.ijoer.com info@ijoer.com

Preface

We would like to present, with great pleasure, the volume-12, Issue-3, March 2026, of a scholarly journal, *International Journal of Engineering Research & Science*. This journal is part of the AD Publications series *in the field of Engineering, Mathematics, Physics, Chemistry and science Research Development*, and is devoted to the gamut of Engineering and Science issues, from theoretical aspects to application-dependent studies and the validation of emerging technologies.

This journal was envisioned and founded to represent the growing needs of Engineering and Science as an emerging and increasingly vital field, now widely recognized as an integral part of scientific and technical investigations. Its mission is to become a voice of the Engineering and Science community, addressing researchers and practitioners in below areas:

Chemical Engineering	
Biomolecular Engineering	Materials Engineering
Molecular Engineering	Process Engineering
Corrosion Engineering	
Civil Engineering	
Environmental Engineering	Geotechnical Engineering
Structural Engineering	Mining Engineering
Transport Engineering	Water resources Engineering
Electrical Engineering	
Power System Engineering	Optical Engineering
Mechanical Engineering	
Acoustical Engineering	Manufacturing Engineering
Optomechanical Engineering	Thermal Engineering
Power plant Engineering	Energy Engineering
Sports Engineering	Vehicle Engineering
Software Engineering	
Computer-aided Engineering	Cryptographic Engineering
Teletraffic Engineering	Web Engineering
System Engineering	
Mathematics	
Arithmetic	Algebra
Number theory	Field theory and polynomials
Analysis	Combinatorics
Geometry and topology	Topology
Probability and Statistics	Computational Science
Physical Science	Operational Research
Physics	
Nuclear and particle physics	Atomic, molecular, and optical physics
Condensed matter physics	Astrophysics
Applied Physics	Modern physics
Philosophy	Core theories

Chemistry	
Analytical chemistry	Biochemistry
Inorganic chemistry	Materials chemistry
Neurochemistry	Nuclear chemistry
Organic chemistry	Physical chemistry
Other Engineering Areas	
Aerospace Engineering	Agricultural Engineering
Applied Engineering	Biomedical Engineering
Biological Engineering	Building services Engineering
Energy Engineering	Railway Engineering
Industrial Engineering	Mechatronics Engineering
Management Engineering	Military Engineering
Petroleum Engineering	Nuclear Engineering
Textile Engineering	Nano Engineering
Algorithm and Computational Complexity	Artificial Intelligence
Electronics & Communication Engineering	Image Processing
Information Retrieval	Low Power VLSI Design
Neural Networks	Plastic Engineering

Each article in this issue provides an example of a concrete industrial application or a case study of the presented methodology to amplify the impact of the contribution. We are very thankful to everybody within that community who supported the idea of creating a new Research with IJOER. We are certain that this issue will be followed by many others, reporting new developments in the Engineering and Science field. This issue would not have been possible without the great support of the Reviewer, Editorial Board members and also with our Advisory Board Members, and we would like to express our sincere thanks to all of them. We would also like to express our gratitude to the editorial staff of AD Publications, who supported us at every stage of the project. It is our hope that this fine collection of articles will be a valuable resource for *IJOER* readers and will stimulate further research into the vibrant area of Engineering and Science Research.

Mukesh Arora

(Chief Editor)

Board Members

Dr. M. Varatha Vijayan

Annauniversity Rank Holder, Commissioned Officer Indian Navy, Ncc Navy Officer (Ex-Serviceman Navy), Best Researcher Awardee, Best Publication Awardee, Tamilnadu Best Innovation & Social Service Awardee From Lions Club.

Dr. Mohamed Abdel Fatah Ashabrawy Moustafa

PhD. in Computer Science - Faculty of Science - Suez Canal University University, 2010, Egypt.

Assistant Professor Computer Science, Prince Sattam bin AbdulAziz University ALkharj, KSA.

Prof.S.Balamurugan

Dr S. Balamurugan is the Head of Research and Development, Quants IS & CS, India. He has authored/co-authored 35 books, 200+ publications in various international journals and conferences and 6 patents to his credit. He was awarded with Three Post-Doctoral Degrees - Doctor of Science (D.Sc.) degree and Two Doctor of Letters (D.Litt) degrees for his significant contribution to research and development in Engineering.

Dr. Mahdi Hosseini

Dr. Mahdi did his Pre-University (12^{th}) in Mathematical Science. Later he received his Bachelor of Engineering with Distinction in Civil Engineering and later he Received both M.Tech. and Ph.D. Degree in Structural Engineering with Grade "A" First Class with Distinction.

Dr. Anil Lamba

Practice Head – Cyber Security, EXL Services Inc., New Jersey USA.

Dr. Anil Lamba is a researcher, an innovator, and an influencer with proven success in spearheading Strategic Information Security Initiatives and Large-scale IT Infrastructure projects across industry verticals. He has helped bring about a profound shift in cybersecurity defense. Throughout his career, he has parlayed his extensive background in security and a deep knowledge to help organizations build and implement strategic cybersecurity solutions. His published researches and conference papers has led to many thought provoking examples for augmenting better security.

Dr. Ali İhsan KAYA

Currently working as Associate Professor in Mehmet Akif Ersoy University, Turkey.

Research Area: Civil Engineering - Building Material - Insulation Materials Applications, Chemistry - Physical Chemistry – Composites.

Dr. Parsa Heydarpour

Ph.D. in Structural Engineering from George Washington University (Jan 2018), GPA=4.00.

Dr. Heba Mahmoud Mohamed Afify

Ph.D degree of philosophy in Biomedical Engineering, Cairo University, Egypt worked as Assistant Professor at MTI University.

Dr. Kalpesh Sunil Kamble (Ph.D., P.Eng., M.Tech, B.E. (Mechanical))

A distinguished academic with a Ph.D. in Mechanical Engineering and 13 Years of extensive teaching and research experience. He is currently a Assistant professor at the SSPM's COE, Kankavli and contributes to several undergraduate and masters programs across Mahrashtra, India.

Dr. Aurora Angela Pisano

Ph.D. in Civil Engineering, Currently Serving as Associate Professor of Solid and Structural Mechanics (scientific discipline area nationally denoted as ICAR/08"–"Scienza delle Costruzioni"), University Mediterranea of Reggio Calabria, Italy.

Dr. Faizullah Mahar

Associate Professor in Department of Electrical Engineering, Balochistan University Engineering & Technology Khuzdar. He is PhD (Electronic Engineering) from IQRA University, Defense View, Karachi, Pakistan.

Prof. Viviane Barrozo da Silva

Graduated in Physics from the Federal University of Paraná (1997), graduated in Electrical Engineering from the Federal University of Rio Grande do Sul - UFRGS (2008), and master's degree in Physics from the Federal University of Rio Grande do Sul (2001).

Dr. S. Kannadhasan

Ph.D (Smart Antennas), M.E (Communication Systems), M.B.A (Human Resources).

Dr. Christo Ananth

Ph.D. Co-operative Networks, M.E. Applied Electronics, B.E Electronics & Communication Engineering Working as Associate Professor, Lecturer and Faculty Advisor/ Department of Electronics & Communication Engineering in Francis Xavier Engineering College, Tirunelveli.

Dr. S.R.Boselin Prabhu

Ph.D, Wireless Sensor Networks, M.E. Network Engineering, Excellent Professional Achievement Award Winner from Society of Professional Engineers Biography Included in Marquis Who's Who in the World (Academic Year 2015 and 2016). Currently Serving as Assistant Professor in the department of ECE in SVS College of Engineering, Coimbatore.

Dr. Balasubramanyam, N

Dr.Balasubramanyam, N working as Faculty in the Department of Mechanical Engineering at S.V.University College of Engineering Tirupati, Andhra Pradesh.

Dr. PAUL P MATHAI

Dr. Paul P Mathai received his Bachelor's degree in Computer Science and Engineering from University of Madras, India. Then he obtained his Master's degree in Computer and Information Technology from Manonmanium Sundaranar University, India. In 2018, he received his Doctor of Philosophy in Computer Science and Engineering from Noorul Islam Centre for Higher Education, Kanyakumari, India.

Dr. M. Ramesh Kumar

Ph.D (Computer Science and Engineering), M.E (Computer Science and Engineering).

Currently working as Associate Professor in VSB College of Engineering Technical Campus, Coimbatore.

Dr. Maheshwar Shrestha

Postdoctoral Research Fellow in DEPT. OF ELE ENGG & COMP SCI, SDSU, Brookings, SD Ph.D, M.Sc. in Electrical Engineering from SOUTH DAKOTA STATE UNIVERSITY, Brookings, SD.

Dr. D. Amaranatha Reddy

Ph.D. (Postdocteral Fellow, Pusan National University, South Korea), M.Sc., B.Sc. : Physics.

Dr. Dibya Prakash Rai

Post Doctoral Fellow (PDF), M.Sc., B.Sc., Working as Assistant Professor in Department of Physics in Pachhuncga University College, Mizoram, India.

Dr. Pankaj Kumar Pal

Ph.D R/S, ECE Deptt., IIT-Roorkee.

Dr. P. Thangam

PhD in Information & Communication Engineering, ME (CSE), BE (Computer Hardware & Software), currently serving as Associate Professor in the Department of Computer Science and Engineering of Coimbatore Institute of Engineering and Technology.

Dr. Pradeep K. Sharma

PhD., M.Phil, M.Sc, B.Sc, in Physics, MBA in System Management, Presently working as Provost and Associate Professor & Head of Department for Physics in University of Engineering & Management, Jaipur.

Dr. R. Devi Priya

Ph.D (CSE), Anna University Chennai in 2013, M.E, B.E (CSE) from Kongu Engineering College, currently working in the Department of Computer Science and Engineering in Kongu Engineering College, Tamil Nadu, India.

Dr. Sandeep

Post-doctoral fellow, Principal Investigator, Young Scientist Scheme Project (DST-SERB), Department of Physics, Mizoram University, Aizawl Mizoram, India- 796001.

Dr. Roberto Volpe

Faculty of Engineering and Architecture, Università degli Studi di Enna "Kore", Cittadella Universitaria, 94100 – Enna (IT).

Dr. S. Kannadhasan

Ph.D (Smart Antennas), M.E (Communication Systems), M.B.A (Human Resources).

Research Area: Engineering Physics, Electromagnetic Field Theory, Electronic Material and Processes, Wireless Communications.

Mr. Bhavinbhai G. Lakhani

An expert in Environmental Technology and Sustainability, with an M.S. from NYIT. Their specialization includes Construction Project Management and Green Building. Currently a Project Controls Specialist Lead at DACK Consulting Solutions, they manage project schedules, resolve delays, and handle claim negotiations. Prior roles as Senior Project Manager at FCS Group and Senior Project Engineer at KUNJ Construction Corp highlight their extensive experience in project estimation, resource management, and on-site supervision.

Mr. Omar Muhammed Neda

Department of Electrical Power Engineering, Sunni Diwan Endowment, Iraq.

Mr. Amit Kumar

Amit Kumar is associated as a Researcher with the Department of Computer Science, College of Information Science and Technology, Nanjing Forestry University, Nanjing, China since 2009. He is working as a State Representative (HP), Spoken Tutorial Project, IIT Bombay promoting and integrating ICT in Literacy through Free and Open Source Software under National Mission on Education through ICT (NMEICT) of MHRD, Govt. of India; in the state of Himachal Pradesh, India.

Mr. Tanvir Singh

Tanvir Singh is acting as Outreach Officer (Punjab and J&K) for MHRD Govt. of India Project: Spoken Tutorial - IIT Bombay fostering IT Literacy through Open Source Technology under National Mission on Education through ICT (NMEICT). He is also acting as Research Associate since 2010 with Nanjing Forestry University, Nanjing, Jiangsu, China in the field of Social and Environmental Sustainability.

Mr. Abilash

MTech in VLSI, BTech in Electronics & Telecommunication engineering through A.M.I.E.T.E from Central Electronics Engineering Research Institute (C.E.E.R.I) Pilani, Industrial Electronics from ATI-EPI Hyderabad, IEEE course in Mechatronics, CSHAM from Birla Institute Of Professional Studies.

Mr. Varun Shukla

M.Tech in ECE from RGPV (Awarded with silver Medal By President of India), Assistant Professor, Dept. of ECE, PSIT, Kanpur.

Mr. Shrikant Harle

Presently working as a Assistant Professor in Civil Engineering field of Prof. Ram Meghe College of Engineering and Management, Amravati. He was Senior Design Engineer (Larsen & Toubro Limited, India).

Mr. Zairi Ismael Rizman

Senior Lecturer, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM) (Terengganu) Malaysia Master (Science) in Microelectronics (2005), Universiti Kebangsaan Malaysia (UKM), Malaysia. Bachelor (Hons.) and Diploma in Electrical Engineering (Communication) (2002), UiTM Shah Alam, Malaysia.

Mr. Ronak

Qualification: M.Tech. in Mechanical Engineering (CAD/CAM), B.E.

Presently working as a Assistant Professor in Mechanical Engineering in ITM Vocational University, Vadodara. Mr. Ronak also worked as Design Engineer at Finstern Engineering Private Limited, Makarpura, Vadodara.

Table of Contents

Volume-12, Issue-3, March 2026

International Journal of Engineering Research & Science (IJOER)
ISSN: [2395-6992]
[Vol-12, Issue-3, March- 2026]
Website: www.ijoer.com
Journal DOI: 10.25125/engineering-journal

Interdependent Networks with Higher-Order Structures

Heng Zhao

Economic and Social Statistics, Guizhou University of Finance and Economics, Guiyang-550025, China
*Corresponding Author

Received: 01 March 2026/ Revised: 08 March 2026/ Accepted: 13 March 2026/ Published: 31-03-2026

Abstract— *In the construction of interdependent networks, the functionality of a group in one layer typically relies on the support of a group in another layer. To investigate the stability of such networks, we propose a framework comprising a bilayer interdependent hypergraph system, where the two layers exhibit mutual dependencies. Our core hypothesis is that the removal of nodes in one layer not only leads to node failures but, more critically, triggers the failure of hyperedges, resulting in iterative cascading failures across layers. Using a bilayer system characterized by a Poisson hyperdegree distribution as an example, we have proven through rigorous analysis how parameter changes affect the robustness of the target network. Overall, our study highlights the critical role of hyperedge interdependence mechanisms and network topological structures in mitigating cascading failures in systems with higher-order interactions, providing valuable insights for the design and optimization of network systems.*

Keywords— Interdependent Networks, Hypergraph, Hyperedge Dependency, Cascading Failure.

I. INTRODUCTION

With the exploration of real-world systems, we find that a single network can no longer meet our real needs[1,2], Networks in the real world usually interact with other networks to varying degrees[3-5] and jointly function to serve the real world. These networks often take the form of multi-layer interdependent networks[6].

In such a network, the cascading failure process is more worthy of study. It usually manifests as the removal of a few nodes triggering iterations through intra-layer failures or inter-layer failures, ultimately bringing the network to a stable state. However, merely the interdependence relationship is far from sufficient. Networks in the real world usually exhibit complex relationships and dependencies, not just pairwise connections. To address this issue, we introduce hypergraphs to reflect high-order interactions and dependencies[7], providing a more comprehensive representation. For example, in social networks[8,9], we can consider the interaction patterns among multiple users, explore the high-order associations among multiple users. In biological networks[10], we can effectively identify the strong and weak relationships between nodes.

Further study the robustness of the network. Percolation theory is a commonly used method to study network robustness[11], played a key role in the study of the collective behavior of systems. There are usually two types of percolation, one of which is bond percolation[12], one is point percolation[13]. So usually, the seepage situation is studied to observe the system behavior of the network. Zhou et al[14] analyzing the characteristics of interdependent networks, a related k-core percolation model is proposed. Given any degree distribution of edge-coupled interdependent networks, the relative sizes of the k-core and the corona cluster as well as the percolation threshold can be obtained through the derivation of self-consistent equations, and then the phase transition of k-core percolation can be analyzed. The proposed k-core percolation model and protection strategy not only help to understand the hierarchical structure of the network, but also provide some guidance for enhancing the resilience of the network against attacks. Secondly, the higher-order interactions in interdependent networks can usually be represented by hyperedges or simplicial complexes. For simplicial complexes, Peng et al[6,15] constructed a two-layer

partial dependence network theoretical model with a simplicial complex, in which failures between nodes occur through the synergistic effects of pairwise interactions and high-order interactions. In this model, removing a node will cause all other nodes in the same simplex to be removed, and due to the dependence between the two networks, node failures will spread through the dependence links between the two networks. This process will occur recursively, ultimately leading to a cascading process. Four types of artificial MIHN models were further constructed, where high-order interactions are still described by simplicial complexes, and inter-layer dependencies are established through one-to-one matching dependence links. The robustness of MIHN was studied by investigating the largest connected component and the percolation threshold. We found that the density of the simplicial complex and the number of network layers affect its percolation behavior.

For hyperedges, Liu et al[16] proposed a percolation model that takes into account the dependence of hyperedges on their internal nodes, and the research reveals the different impacts of the hyperdegree distribution on the system robustness in single-layer and double-layer hypergraphs. In the real world, not only do nodes have interdependent relationships, but edges do as well. Qian et al[17] proposed an interdependent hypergraph model considering the inter-layer node dependency. The cascading failures in hypergraphs with different inter-layer dependencies were studied. The maximum attack intensity that the network can withstand was determined through theoretical analysis, as well as how its robustness changes under different attack intensities. The multi-layer hypergraph can represent the relationships between nodes more clearly. However, there is less research on identifying important nodes within this framework. Wang et al[18] proposed a method named HCT to fill this gap. The global centrality of nodes in the entire network can be calculated. Compared with other methods, the important nodes identified by HCT exhibit stronger propagation capabilities, and removing these nodes will seriously damage the connectivity and robustness of the network. To further reveal the profound impact of group support on the resilience of the system against cascading failures, Chen et al [19] designed a framework consisting of a two-layer interdependent hypergraph system, where the nodes in one layer can obtain support through the hyperedges in the other layer. The article derived the critical threshold of the initial node survival probability that marks the second-order phase transition point.

There is also an understanding of network resilience. Network resilience measures the degree of performance degradation of a network after being perturbed and its recovery ability, and is closely related to the ability to resist cascading failures. Li et al[20] proposed three resilience enhancement strategies based on the node capacity redundancy at different structural scales, and developed a network resilience evaluation method that considers both the structure and node load. The performance of the enhancement strategies is closely related to the node capacity redundancy. Specifically, when enhancing nodes with larger capacity redundancy, the enhancement efficiency is higher. In addition, the heterogeneity of node load has a profound impact on the enhancement efficiency. Lv et al [21] proposed a resilience assessment model to predict the performance of interdependent networks against cascading failures. This model can accurately monitor the activities of each node during the cascading process.

In the real world, systems often exist in the form of groups, so that the network not only needs to consider the intra-layer dependencies but also the inter-layer dependencies. Furthermore, it is necessary to consider whether this dependence is node dependence or hyperedge dependence. Therefore, we will combine these two aspects and use hypergraphs to study the robustness of two-layer partially interdependent networks.

II. MODEL AND THEORY

2.1 Model:

We construct a bilayer hypergraph system with hyperedge interdependencies, denoted as layers A and B. Layer A consists of N_A nodes and M_A hyperedges, while layer B comprises N_B nodes and M_B hyperedges. The hyperdegree of each node, denoted by k, represents the number of hyperedges to which the node belongs. The hyperdegree distributions of layers A and B follow $P_A(k)$ and $P_B(k)$, respectively. Similarly, the cardinality of a hyperedge, denoted by m, indicates the number of nodes contained in each hyperedge, with distributions following $Q_A(m)$ and $Q_B(m)$, respectively. Regarding interlayer dependencies, we define that hyperedges in layers A and B are interdependent with probability λ.

Due to the interdependence between network systems being maintained through hyperedges, the cascading failure process involves both node failures and hyperedge failures. First, we consider node failures. In layer A, nodes are initially present

with probability p (equivalently, removed with probability $1-p$), where p represents the initial node survival probability. For a node to remain functional, it must satisfy the following conditions: (1) the node must belong to the Giant Connected Component (GCC) of its layer to ensure intralayer connectivity; (2) all hyperedges to which the node belongs must remain functional.

Second, we address hyperedge failures. Each hyperedge initially exists in a functional state. To remain functional after node failures, a hyperedge must satisfy the following conditions: (1) following node failures, the functional state of the hyperedge is reassessed; we define a tolerance coefficient $\alpha \in [0,1]$ (typically taken as $\alpha = 0.5$ in this study), and a hyperedge fails if the proportion of its constituent nodes that are functional falls below α; (2) if a hyperedge's interdependent partner hyperedge in the opposite layer fails, the hyperedge itself fails.

This cascading failure process, triggered by initial node removal, iterates between the two layers until no additional node or hyperedge failures occur. The final sizes of the Giant Connected Components in layers A and B are denoted as G_A and G_B, respectively.

The cascading failure process in the model is illustrated in Figure 1. (a) Initially, node 4 in layer A is removed. (b) Nodes (1, 2, 3) fail because they become disconnected from the GCC, and hyperedges e_1^A and e_2^A fail as their functional node proportions fall below $\alpha = 0.5$. (c) Due to the interdependence between layers A and B, hyperedges e_1^B and e_2^B in layer B fail, and hyperedge e_3^B fails because its functional node proportion equals exactly 0.5. (d) Layer B further affects layer A, causing hyperedge e_3^A in layer A to fail, at which point the system reaches a stable state.

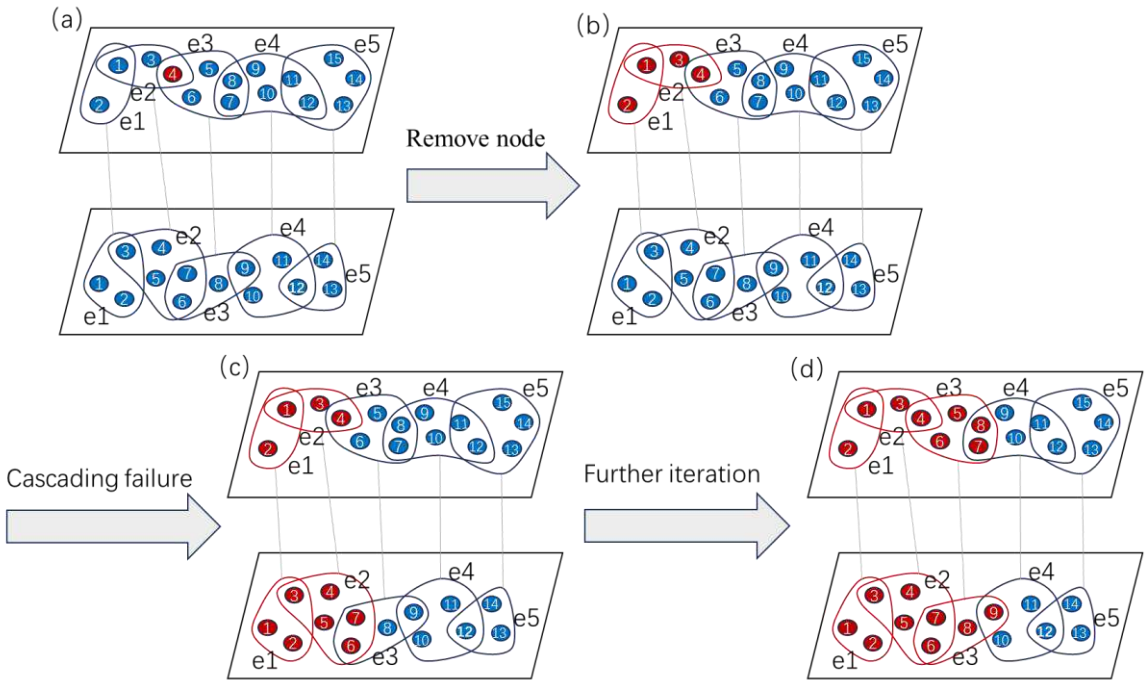


FIGURE 1: Schematic diagram of the cascading failure process in bilayer hypergraphs A and B. The connections between the two layers represent the interdependence established between a hyperedge in one layer and its corresponding dependent hyperedge in the other layer. Blue nodes indicate functional nodes, black hyperedges represent functional hyperedges, red nodes denote failed nodes, and red hyperedges indicate hyperedges that have failed due to degraded functionality.

2.2 Theory

To characterize the structure of the two-layer network and the interdependence between them, we introduce generating functions to simplify the key information regarding the hyperdegree distribution and cardinality distribution in each layer[19].

Since this study considers a bilayer network, the generating functions for the hyperdegree distributions of networks A and B are similar and are both functions of k:

$$\begin{cases} G_{k0}^{A}(x) = \sum_{k=0}^{\infty} P^{A}(k) x^{k} \\ G_{k0}^{B}(x) = \sum_{k=0}^{\infty} P^{B}(k) x^{k} \end{cases} \tag{1}$$

The generating function for the distribution of excess hyperdegree is similarly expressed as:

$$\begin{cases} G_{k1}^{A}(x) = \sum_{k=1}^{\infty} \frac{kP^{A}(k)}{\langle k \rangle} x^{k-1} = G_{k0}^{A'}(x) / G_{k0}^{A'}(1) \\ G_{k1}^{B}(x) = \sum_{k=1}^{\infty} \frac{kP^{B}(k)}{\langle k \rangle} x^{k-1} = G_{k0}^{B'}(x) / G_{k0}^{B'}(1) \end{cases} \tag{2}$$

The generating function for the cardinality distribution is defined as:

$$\begin{cases} G_{m0}^{A}(x) = \sum_{m=0}^{\infty} Q^{A}(m) x^{m} \\ G_{m0}^{B}(x) = \sum_{m=0}^{\infty} Q^{B}(m) x^{m} \end{cases} \tag{3}$$

The generating function for the distribution of excess bases is defined as:

$$\begin{cases} G_{m1}^{A}(x) = \sum_{m=1}^{\infty} \frac{mQ^{A}(m)}{\langle m \rangle} x^{m-1} = G_{m0}^{A'}(x) / G_{m0}^{A'}(1) \\ G_{m1}^{B}(x) = \sum_{m=1}^{\infty} \frac{mQ^{B}(m)}{\langle m \rangle} x^{m-1} = G_{m0}^{B'}(x) / G_{m0}^{B'}(1) \end{cases} \tag{4}$$

As shown in Figure 1, the cascading failure process unfolds in a discrete iterative manner, progressing from stage $n=1$ to $n=\infty$ in both layers of the network.

We denote the probabilities of hyperedges functioning normally in layers A and B at the n stage as T_n^A and T_n^B, respectively. Correspondingly, the probability that a dependent hyperedge in layer $A(B)$ lacks support from normally functioning hyperedges in layer $B(A)$ at the n stage is given by the following expression:

$$\begin{cases} u_n^A = q\left(1 - T_{n-1}^B\right) \\ u_n^B = q\left(1 - T_{n-1}^A\right) \end{cases} \tag{5}$$

The proportion of nodes that maintain their functions in layer $A(B)$ of the nth stage can be expressed as:

$$\begin{cases} p_n^A = r^A\left(1 - qu_n^A\right) \\ p_n^B = r^B\left(1 - qu_n^B\right) \end{cases} \tag{6}$$

T_n^A and T_n^B can be expressed as:

$$\begin{cases} T_n^A = \sum_{m=0}^{\infty} Q^A(m) \sum_{j=0}^{m} \binom{m}{j} \left\{1-\left[G_{k1}^A\left(1-f_n^A\right)\right]^{m-j}\right\}\left(p_n^A\right)^{m-j}\left(1-p_n^A\right)^j I\left(\frac{m-j}{m}>\delta\right) \\ T_n^B = \sum_{m=0}^{\infty} Q^B(m) \sum_{j=0}^{m} \binom{m}{j} \left\{1-\left[G_{k1}^B\left(1-f_n^B\right)\right]^{m-j}\right\}\left(p_n^B\right)^{m-j}\left(1-p_n^B\right)^j I\left(\frac{m-j}{m}>\delta\right) \end{cases} \tag{7}$$

Among them, f_n^A and f_n^B represents the probability that a randomly selected hyper-edge through a random node can be connected to the giant connected components (GCC) of each layer. First, select a random hyperedge with cardinality m according to the distribution $Q^A(m)$ or $Q^B(m)$ .Then, $\binom{m}{j}$ represents the situation where there are j failed nodes among the m nodes of the selected hyperedge. $1-\left[G_{k1}^A\left(1-f_n^A\right)\right]^{m-1-j}$ or $1-\left[G_{k1}^B\left(1-f_n^B\right)\right]^{m-1-j}$ represents the probability that at least one of the $m-j$ functional nodes on the remaining hyperedges can be connected to the Giant Connected Component (GCC) in a random hyperedge with a base of m . $I(\cdot)$ is an indicator function. When the condition in the parentheses holds, it takes the value of 1; otherwise, it takes the value of 0. It is used to determine whether the hyperedge state exceeds the tolerance coefficient δ . Generally, the tolerance coefficient δ is first taken as 0.5.

f_n^A and f_n^B can be expressed as:

$$\begin{cases} f_n^A = \sum_{m=1}^{\infty} \frac{mQ^A(m)}{\langle m\rangle} \sum_{j=0}^{m-1} \binom{m-1}{j} \left\{1-\left[G_{k1}^A\left(1-f_n^A\right)\right]^{m-1-j}\right\}\left(p_n^A\right)^{m-1-j}\left(1-p_n^A\right)^j I\left(\frac{m-1-j}{m-1}>\delta\right) \\ f_n^B = \sum_{m=1}^{\infty} \frac{mQ^B(m)}{\langle m\rangle} \sum_{j=0}^{m-1} \binom{m-1}{j} \left\{1-\left[G_{k1}^B\left(1-f_n^B\right)\right]^{m-1-j}\right\}\left(p_n^B\right)^{m-1-j}\left(1-p_n^B\right)^j I\left(\frac{m-1-j}{m-1}>\delta\right) \end{cases} \tag{8}$$

Therefore, the proportion of nodes S_n^A and S_n^B in the Giant Connected Component (GCC) can be obtained from the following formula:

$$\begin{cases} S_n^A = p_n^A\left[1-G_{k0}^A\left(1-f_n^A\right)\right] \\ S_n^B = p_n^B\left[1-G_{k0}^B\left(1-f_n^B\right)\right] \end{cases} \tag{9}$$

When the cascading failure process terminates, they converge to their steady state values respectively: $p_\infty^A, p_\infty^B, T_\infty^A, T_\infty^B, f_\infty^A, f_\infty^B, S_\infty^A, S_\infty^B$. All the above formulas can form a self-consistent system of equations, written in the form of a generating function:

$$\begin{cases} T_\infty^A = I\left(\frac{m-j}{m}>\delta\right)\left(1-G_{m0}^A\left(1-p_\infty^A+p_\infty^A G_{k1}^A\left(1-f_\infty^A\right)\right)\right) \\ f_\infty^A = I\left(\frac{m-j}{m}>\delta\right)\left(1-G_{m1}^A\left(1-p_\infty^A+p_\infty^A G_{k1}^A\left(1-f_\infty^A\right)\right)\right) \end{cases} \tag{10}$$

$$\begin{cases} T_{\infty}^{B} = I\left(\dfrac{m-j}{m} > \delta\right)\left(1 - G_{m0}^{B}\left(1 - p_{\infty}^{B} + p_{\infty}^{B} G_{k1}^{B}\left(1 - f_{\infty}^{B}\right)\right)\right) \\ f_{\infty}^{B} = I\left(\dfrac{m-j}{m} > \delta\right)\left(1 - G_{m1}^{B}\left(1 - p_{\infty}^{B} + p_{\infty}^{B} G_{k1}^{B}\left(1 - f_{\infty}^{B}\right)\right)\right) \end{cases} \tag{11}$$

$$\begin{cases} S_{\infty}^{A} = p_{\infty}^{A}\left[1 - G_{k0}^{A}\left(1 - f_{\infty}^{A}\right)\right] \\ S_{\infty}^{B} = p_{\infty}^{B}\left[1 - G_{k0}^{B}\left(1 - f_{\infty}^{B}\right)\right] \end{cases} \tag{12}$$

First, assume that both layers follow a Poisson cardinality distribution: $e^{-<m>} < m >^{m} / m!$, This simplifies the solution of the equation. Therefore, we have:

$$\begin{cases} G_{m0}^{A}(x) = G_{m1}^{A}(x) \\ G_{m0}^{B}(x) = G_{m1}^{B}(x) \end{cases} \tag{13}$$

$$\begin{cases} f_{\infty}^{A} = T_{\infty}^{A} \\ f_{\infty}^{B} = T_{\infty}^{B} \end{cases} \tag{14}$$

$$\begin{cases} f_{\infty}^{A} = I\left(\dfrac{m-j}{m} > \delta\right)\left(1 - G_{m0}^{A}\left(1 - p_{\infty}^{A} + p_{\infty}^{A} G_{k1}^{A}\left(1 - f_{\infty}^{A}\right)\right)\right) \\ f_{\infty}^{B} = I\left(\dfrac{m-j}{m} > \delta\right)\left(1 - G_{m0}^{B}\left(1 - p_{\infty}^{B} + p_{\infty}^{B} G_{k1}^{B}\left(1 - f_{\infty}^{B}\right)\right)\right) \end{cases} \tag{15}$$

Furthermore, if the two hypergraph layers have identical cardinality and hyperdegree distributions, we can derive:

$$\begin{cases} G_{m0}^{A}(x) = G_{m0}^{B}(x) = G_{m0}(x) \\ G_{k0}^{A}(x) = G_{k0}^{B}(x) = G_{k0}(x) \end{cases} \tag{16}$$

If $q_A = q_B = q, r^A = r^B = r$, we can derive:

$$f_{\infty}^{A} = f_{\infty}^{B} = f \tag{17}$$

$$f = I(\frac{m-j}{m} > \delta)\left(1 - G_{m0}\left(1 - r\left[1 - q^{2}(1-f)\right]\left[1 - G_{k1}(1-f)\right]\right)\right) \tag{18}$$

$$S = r\left[1 - q^{2}(1-f)\right]\left[1 - G_{k0}(1-f)\right] \tag{19}$$

III. RESULTS AND DISCUSSION

In Figure 2, we can observe that, under the condition of fixing the average hyperdegree, the average hyperedge cardinality, and the tolerance coefficient, as the proportion of node removal increases, the change of the giant connected component (GCC). For the overall trend, all curves gradually decrease as $1-r$ increases and then tend to 0, which conforms to the general law when the network is under attack. The smaller the q is, the size of GCC can still maintain a relatively high proportion for a period of time; conversely, the larger the q is, the more sensitive the network is to node removal, and a

small-scale node damage will lead to network collapse. As q decreases, the network characteristics tend to transform into a more continuous phase transition.

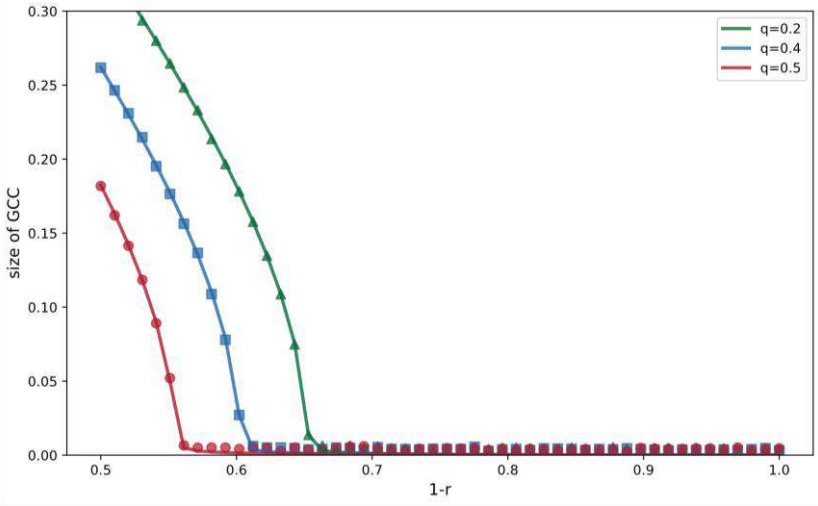


FIGURE 2: The trend graph of the GCC size changing with the node removal ratio. Different colors represent the change of the GCC size with the node removal ratio under different dependency ratios.

From Figure 3, we can see that as the dependency ratio increases, the network is more likely to collapse. Moreover, the larger the average hyperdegree, or the average hyperedge cardinality, the stronger the network's invulnerability, and the less likely it is to collapse as increases. From the figure, we can further learn that, excluding the situations where the dependency ratio, average hyperdegree, and average hyperedge cardinality are relatively small, as the dependency ratio increases and the average hyperdegree, and average hyperedge cardinality become larger, the network characteristics tend to transform into a more continuous phase transition.

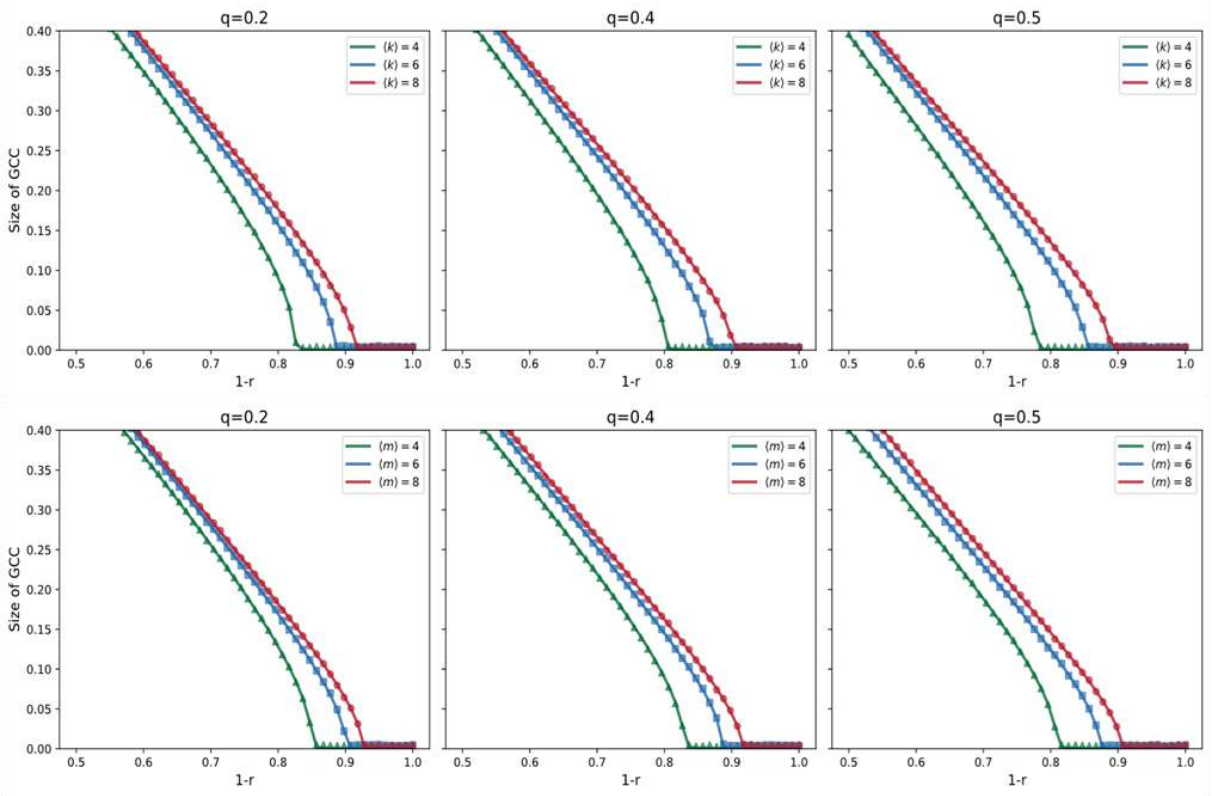


FIGURE 3: It shows the variation of the GCC size with the increase of the removal proportion under different parameter settings, that is, under different average hyperdegree and different average hyperedge cardinalities.

The first sub-figure in Figure 4 illustrates the variation of the size of the Giant Connected Component (GCC) under different tolerance coefficients as the proportion of removed nodes increases. Under a low tolerance coefficient, when the initial failure proportion is relatively small, the size of the GCC has already started to decrease significantly; as gradually increases, the GCC approaches 0. When the tolerance coefficient is small, the network is extremely sensitive to the initial failure, and the critical value of the phase transition is low. As the tolerance coefficient becomes larger and larger, a larger initial failure proportion is required for the size of the GCC to decrease, which indicates that the network can withstand a high-proportion initial failure, and the critical value of the phase transition is high, that is, only a strong disturbance can trigger the collapse of the network.

The second sub-figure shows the variation of the size of the GCC under the combined effect of the node retention proportion and the tolerance coefficient. It can be known from the color bar that the redder the color, the larger the size of the GCC. Fixing, it can be seen that a larger is better; fixing, it can be seen that a larger is better. Therefore, only when both and become larger and larger will the size of the GCC become larger and larger. The white lines in the figure represent contour lines, which means that for any combination of and on the line segment, the size of the GCC is constant.

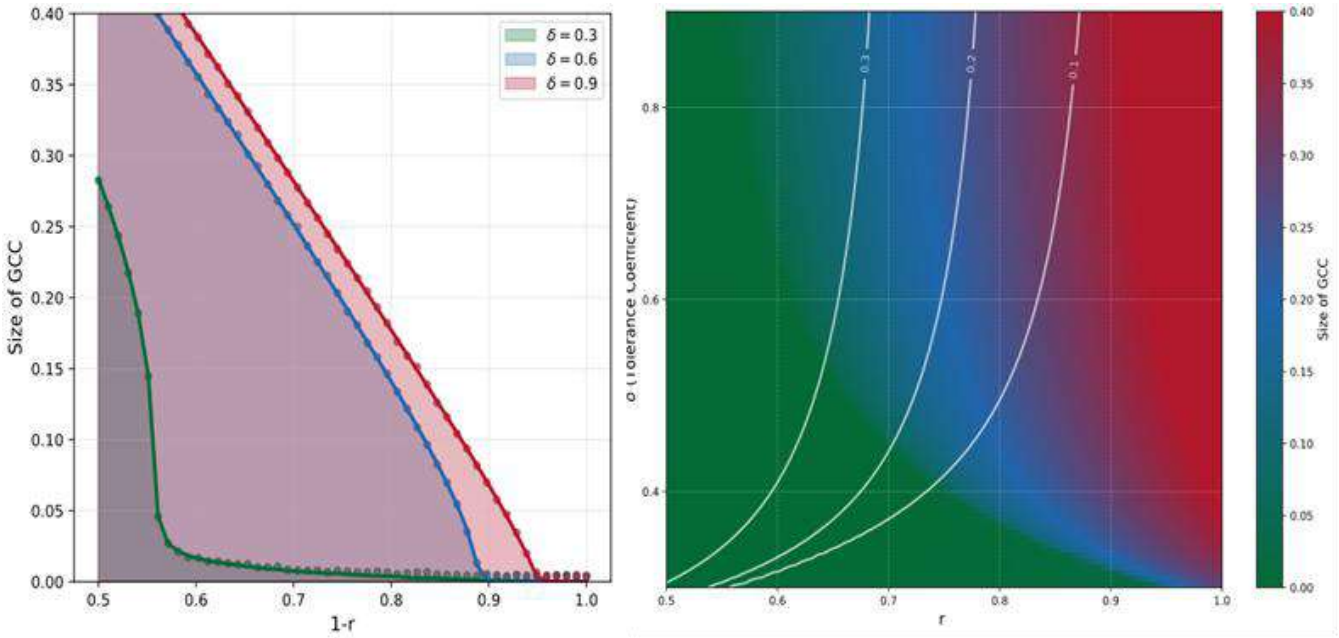


FIGURE 4: Respectively show the changes in the size of GCC under different tolerance coefficient conditions as the dependency proportion increases, and the changes in the size of GCC under the combined effect of the node retention proportion and the tolerance coefficient.

From Figure 5, the presented results demonstrate the variations in the size of GCC under different dependency ratios and various parameter combinations. Firstly, a larger dependency ratio renders the network more prone to collapse. Secondly, with other parameters kept constant, increasing the tolerance coefficient emerges as the optimal approach to enhance the network's invulnerability. When other parameters are kept constant and only the average hyperdegree or the average hyperedge cardinality is changed, the phase transition of the network will change. The larger and are, the slower the GCC decreases with and the network can withstand a higher proportion of node removal. In addition, in the case of a relatively small dependency ratio, increasing the average hyperedge cardinality can enhance the invulnerability of the network more effectively than increasing the average hyperdegree.

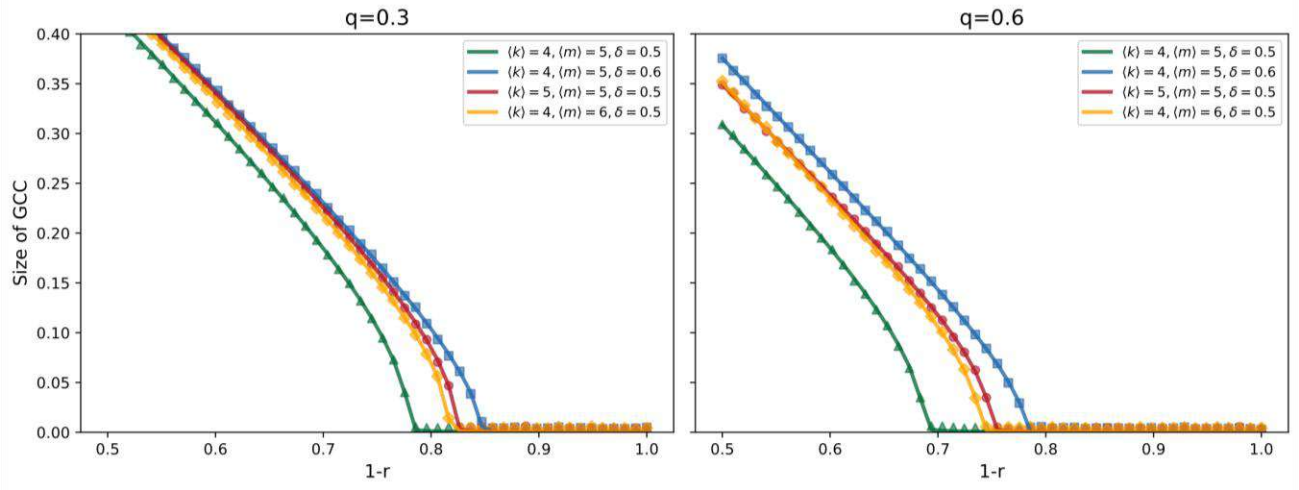


FIGURE 5: Results of parameter combinations under different dependency ratio conditions.

IV. CONCLUSION

This study examines cascading failures and phase transitions in a bilayer hypergraph system with hyperedge interdependencies. Key findings include: cascading failures arise from coupled node and hyperedge failures, driven by hyperedges dependence on both constituent node functionality and interdependent hyperedges across layers, iterating until a steady state. The giant connected component (GCC) exhibits phase transitions with increasing initial node removal. Lower dependency ratios or larger delay critical transitions, and makes the phase transition exhibit more continuous characteristics. Higher or smaller heighten sensitivity to initial failures, lowering collapse thresholds. Node retention proportion and synergistically affect GCC size: larger values of both preserve GCC integrity, with contour lines indicating (,) combinations yielding equivalent GCC sizes. Larger average hyperdegree and cardinality enhance invulnerability, mitigating collapse under higher. With other parameters remaining unchanged, increasing the tolerance coefficient is the optimal way to improve the network's invulnerability, In the case of a relatively small dependency ratio, increasing the average hyperedge cardinality yields better results than increasing the average hyperdegree. These results highlight hyperedge interdependencies and topology in bilayer hypergraph resilience, aiding network optimization against cascading failures.

ACKNOWLEDGEMENT

The authors have no acknowledgements to declare.

V. CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*(4-5), 175–308.

[2] Johnston, W. J., Peters, L. D., & Gassenheimer, J. (2006). Questions about network dynamics: Characteristics, structures, and interactions. *Journal of Business Research*, *59*(8), 945–954.

[3] Du, W.-B., Zhou, X.-L., Lordan, O., Wang, Z., Zhao, C., & Zhu, Y.-B. (2016). Analysis of the Chinese Airline Network as multi-layer networks. *Transportation Research Part E: Logistics and Transportation Review*, *89*, 108–116.

[4] Zhang, L., Du, H., Zhao, Y., & Wu, D. (2019). Drawing topological properties from a multi-layered network: The case of an air transport network in "the Belt and Road" region. *Habitat International*, *93*, Article 102044.

[5] Jia, C. X., & Liu, R. R. (2025). Cascading dynamics in double-layer hypergraphs with higher-order inter-layer interdependencies. *Reliability Engineering & System Safety*, *255*, Article 110678.

[6] Peng, H., Zhao, Y., Zhao, D., & Zhong, M. (2025). Robustness of multilayer interdependent higher-order network. *Journal of Network and Computer Applications*, *235*, Article 104067.

[7] Duan, Y., Huang, J., Deng, H., & Liu, Z. (2024). Robustness of hypergraph under attack with limited information based on percolation theory. *Chaos, Solitons & Fractals*, *180*, Article 114523.

[8] Yang, W., Wang, G., Bhuiyan, M. Z. A., & Choo, K.-K. R. (2017). Hypergraph partitioning for social networks based on information entropy modularity. *Journal of Network and Computer Applications*, *86*, 59–71.

[9] Wu, D., Tang, M., Zhang, S., & Li, X. (2025). HMKRec: Optimize multi-user representation by hypergraph motifs for knowledge-aware recommendation. *Engineering Applications of Artificial Intelligence, 139*, Article 109567.

[10] Gopalakrishnan, S., Sridharan, S., Nayak, S. R., & Swaminathan, P. (2022). Central hubs prediction for bio networks by directed hypergraph - GA with validation to COVID-19 PPI. *Pattern Recognition Letters, 153*, 154–162.

[11] Li, M., Liu, R. R., Lü, L., Hu, M.-B., Xu, S., & Zhang, Y.-C. (2021). Percolation on complex networks: Theory and application. *Physics Reports, 907*, 1–68.

[12] Hackett, A., Cellai, D., Gómez, S., Arenas, A., & Gleeson, J. P. (2016). Bond percolation on multiplex networks. *Physical Review X, 6*(2), Article 021002.

[13] Diskin, S., & Krivelevich, M. (2023). Site percolation on pseudo-random graphs. *Random Structures & Algorithms, 62*(3), 567–595.

[14] Zhou, L., Liao, H., Tan, F., & Chen, X. (2025). K-core percolation and node protection strategy for edge-coupling partially interdependent networks. *Expert Systems with Applications, 260*, Article 125456.

[15] Peng, H., Zhao, Y., Zhao, D., & Zhong, M. (2023). Robustness of higher-order interdependent networks. *Chaos, Solitons & Fractals, 172*, Article 113567.

[16] Liu, R. R., Chu, C., & Meng, F. (2023). Higher-order interdependent percolation on hypergraphs. *Chaos, Solitons & Fractals, 174*, Article 113789.

[17] Qian, C., Zhao, D., Zhong, M., & Peng, H. (2024). Cascading failures on interdependent hypergraph. *Communications in Nonlinear Science and Numerical Simulation, 128*, Article 107678.

[18] Wang, P., Ling, G., Zhao, P., & Li, Y. (2024). Identification of important nodes in multi-layer hypergraphs based on fuzzy gravity model and node centrality distribution characteristics. *Chaos, Solitons & Fractals, 180*, Article 114567.

[19] Chen, L., Lu, J., Wang, Y., & Zhang, X. (2025). Cascading failures with group support in interdependent hypergraphs. *Chaos, Solitons & Fractals, 190*, Article 115678.

[20] Li, J., Wang, Y., Zhong, J., & Chen, G. (2022). Network resilience assessment and reinforcement strategy against cascading failure. *Chaos, Solitons & Fractals, 160*, Article 112234.

[21] Lv, C., Lei, Y., Zhang, Y., & Wang, H. (2025). Resilience of the interdependent network against cascade failure. *Chaos, Solitons & Fractals, 192*, Article 115901.

International Journal of Engineering Research & Science (IJOER)
ISSN: [2395-6992]
[Vol-12, Issue-3, March- 2026]
Website: www.ijoer.com
Journal DOI: 10.25125/engineering-journal

The Impact of Implicit Guarantees and Market-Based Guarantees on the Issuance Spreads of Urban Investment Bonds

Qiu Liu

Economic and Social Statistics, Guizhou University of Finance and Economics, Guiyang-550025, China

Received: 03 March 2026/ Revised: 11 March 2026/ Accepted: 17 March 2026/ Published: 31-03-2026

Abstract— *Preventing and defusing local government debt risks is a core issue in current macroeconomic governance. Against the backdrop of policies that promote the separation of credit between urban investment platforms and local governments, market pricing remains constrained by expectations of implicit guarantees. Using urban investment bonds issued between 2014 and 2024 as the research sample, this paper constructs an issuer-bond and guarantor-bond bipartite network, introduces two dynamic indicators: issuer node degree and guarantor node degree, and employs a quasi-natural experiment based on a series of debt resolution policies to examine the interactive effects of policy regulation and network structure on issuance spreads. The empirical results show that policy effects exhibit significant heterogeneity; the pricing weight of network characteristics is reshaped with the policy environment, where issuer node degree transforms into a market-oriented signal, while guarantor node degree reverses into a hub for risk transmission. Urban investment bond pricing deviates systematically from traditional financial theories, with variables capturing government credit linkages acting as the core pricing determinants. The findings of this paper indicate that the risk pricing of urban investment bonds is undergoing a transition from relationship dependence to fundamental risk pricing, which provides empirical support for identifying implicit guarantees and preventing risk contagion within guarantee networks.*

Keywords— Urban Investment Bonds (UIBs); Issuance Spread; Implicit Guarantee; Network Analysis.

I. INTRODUCTION

1.1 Research Background and Significance:

Against the backdrop of the fiscal revenue-expenditure gap caused by the 1994 tax-sharing reform, urban investment bonds emerged as a "quasi-municipal bond" with Chinese characteristics. By the end of 2024, the outstanding balance of implicit local government debt had reached 10.5 trillion yuan, and its risk transmission has become a core issue in the establishment of a joint prevention and control mechanism for fiscal and financial risks (Qiu et al., 2022). Since the issuance of Document No. 43 [2014] of the State Council, which established the debt governance framework, China has continuously promoted the resolution of implicit debt through Document No. 50, Document No. 35, and the "package debt resolution" policies. However, discrepancies still exist between the market pricing mechanism and policy orientation, highlighting the path-dependent nature of implicit guarantee expectations (Cao, 2023). Notably, existing studies are mostly limited to the analysis of traditional financial indicators and fail to reveal how the structural position of issuers in the debt network affects risk pricing. Against the background that policies continue to weaken unified implicit guarantees, structural indicators such as network node degree may reshape the formation mechanism of credit spreads through dual channels of information transmission and risk exposure. By constructing two dynamic network indicators—issuer node degree and guarantor node degree—this paper systematically analyzes the interactive effects of policy regulation and network structure on the issuance spreads of urban investment bonds, aiming to provide theoretical references and practical evidence for improving the debt governance system and preventing systemic financial risks.

1.2 Literature Review:

Domestic and foreign scholars have established a multi-level research system on the pricing mechanism of bonds backed by government credit. At the macroeconomic level, early studies constructed an analytical framework including variables such as money supply and GDP (Altman, 1990). Subsequent research verified the transmission effects of the economic cycle and risk-free interest rate on credit spreads (Collin-Dufresne et al., 2001). Chinese scholars have echoed this perspective, finding that financial environment optimization (Pan et al., 2015), fiscal sustainability (Yan & Ma, 2024) can reduce issuance costs, while risks of urban investment bonds may push up treasury bond premiums and exert systemic impacts (Niu et al., 2016).

At the micro level, bond maturity (Fons, 1994), embedded options (Kalotay, 1997), and guarantee arrangements (Zhang & Jiao, 2017) serve as core pricing determinants. Institutional factors represent a key dimension highlighting Chinese characteristics. Foreign studies emphasize the roles of fiscal transparency and governance structure (Bo et al., 2023; Feld, 2017). Domestic research has deeply revealed the dominant role of implicit government guarantees and their substitution effect with market-oriented guarantees (Ouyang & Wang, 2024; Chen et al., 2024), while also noting the upward pressure of administrative intervention on financing costs (Zheng et al., 2025).

Regional economies and issuer characteristics form the micro-foundation of pricing. Pairing assistance (Liang et al., 2025) and equity connections (Tong et al., 2024) can improve financing conditions. In recent years, studies have begun to regard local government economic activities as a complex network system, confirming that risk contagion and information spillover embedded in network structure profoundly affect credit pricing (Mao et al., 2024; Wu et al., 2021).

Nevertheless, existing literature still has obvious limitations. First, most studies treat guarantees as isolated dummy variables, ignoring the impact of the systematic network formed by issuer-guarantor relationships and its topological characteristics on risk pricing. Second, they fail to dynamically reveal the structural evolution of guarantee networks across different policy cycles and their moderating effects on bond pricing logic. Against this background, this paper adopts a network analysis perspective. By constructing two dynamic indicators—issuer node degree and guarantor node degree—we systematically examine the impact of network characteristics on the issuance spreads of urban investment bonds. This study aims to expand the analytical boundaries of bond pricing theory and provide implications for preventing systemic risks.

II. CURRENT SITUATION AND THEORETICAL ANALYSIS

2.1 Analysis of the Current Situation of Urban Investment Bonds:

China’s urban investment bonds originated in Shanghai. On July 22, 1992, with the approval and authorization of the Shanghai Municipal Government to broaden financing channels for urban construction, Shanghai established a specialized investment and holding company dedicated to raising and managing funds for urban construction and maintenance: Shanghai Urban Construction Investment & Development Corporation.

On April 15, 1993, the company issued its first urban construction bond worth 500 million yuan with a maturity of 2 years and a coupon rate of 10.5%. This became China’s first urban investment bond, opening a new chapter in the development of China’s bond market.

2.1.1 Bond Maturity:

As shown in Table 1, in terms of the evolution of the term structure, China’s urban investment bonds exhibit a dual trend of coexisting short-termization and medium- to long-termization, reflecting the complex interaction between market liquidity preferences and policy guidance. The term structure of urban investment bonds underwent a landmark shift in 2024, characterized by a move from the previous concentrated pattern dominated by 3-year bonds to a more extended distribution led by 5-year bonds. This change is not a simple maturity substitution: the share of 5-year bonds surged to 54.5%, while that of 3-year bonds dropped from a high of 71.6% in 2023 to 36.8%, indicating a profound adjustment in market logic. Against the policy backdrop of “controlling new debt and digesting existing debt”, this signals that urban investment platforms are proactively managing their debt maturity structure and pursuing long-term stability. It also reflects that investors, guided by policies, have granted longer risk exposure to platforms with relatively sound fundamentals. Although short-term bonds still account for about 10% to meet liquidity needs and ultra-long-term bonds remain stable, the noticeable upward shift in the maturity center is quietly reshaping the risk pricing and rollover model of the urban investment bond market.

TABLE 1
BOND MATURITY OF URBAN INVESTMENT BONDS

Maturity	2020	2021	2022	2023	2024
<1	761	237	1036	146	833
1	379	85	426	108	551
3	1157	352	1222	1753	3083
5	1891	1001	1176	1241	4561
7	948	746	661	430	220
>=10	138	122	49	51	211

Data source: Wind Database, Resset Database. The same below.

2.1.2 Analysis of the Guarantee Network Structure of Urban Investment Bonds:

The visual map of the issuer-guarantor network intuitively presents the coexistence of high fragmentation and local centralization in China's local debt guarantee system. As shown in Figure 1, the nationwide guarantee network of outstanding urban investment bonds at the end of 2024 contains 1,750 nodes with a modularity coefficient as high as 0.882, forming 301 independent communities. This indicates that the so-called national guarantee network is actually a complex composed of hundreds of locally connected, highly isolated local credit clusters.

Network parameters further confirm this feature: the average degree is only 0.994 and the graph density is merely 0.001. Guarantee relationships mostly present a star structure radiating from a single core node, with credit resources highly concentrated in provincial guarantee groups or key municipal platforms. For instance, Jiangsu Credit Re-guarantee Group, Chongqing Three Gorges Financing Guarantee, and Tianfu Credit Enhancement are all core nodes of local clusters.

This structure reveals that, at the macro level, credit resources are fragmented due to administrative and regional barriers, lacking a cross-regional risk-sharing mechanism. At the micro level, each sub-network relies on a small number of core nodes for internal control and coordination. Even under policy interventions such as the "package debt resolution" plan, the basic logic of the guarantee network has not been fundamentally restructured, and the boundaries of local implicit guarantees remain clear. The stability of the entire system highly depends on the individual robustness of local core nodes, and its systemic resilience deserves continuous attention.

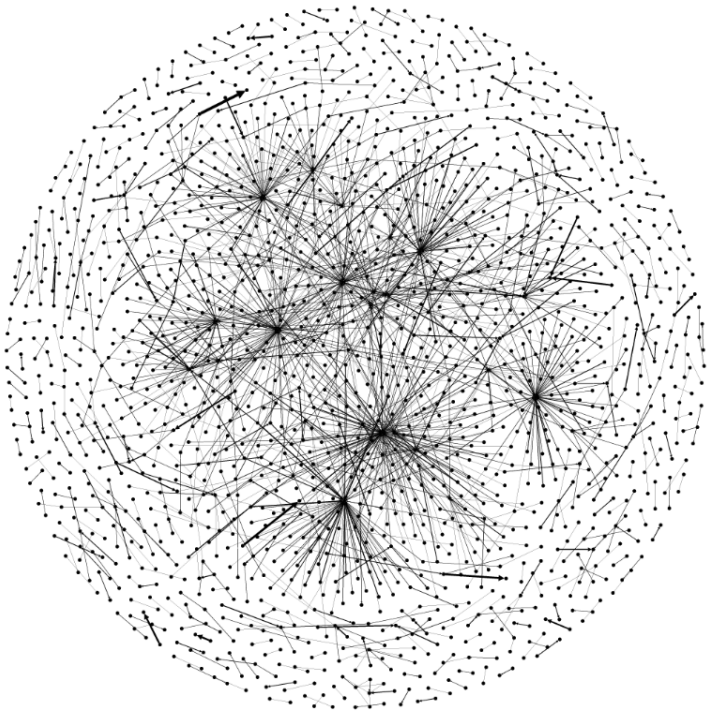


FIGURE 1: Guarantee Network of Outstanding Urban Investment Bonds in China by the End of 2024

2.2 Theoretical Analysis and Research Hypotheses:

2.2.1 Analysis of Implicit Guarantees:

Most financing platforms are established with authorization from local governments, and their debt obligations often carry implicit government credit backing. The market generally expects the public sector to intervene and provide bailouts in the event of default, which significantly depresses the risk pricing of urban investment bonds (Wang et al., 2016), but also distorts risk signals and misallocates financial resources (Luo & Liu, 2016).

To alleviate such distortions, China has successively introduced a series of debt resolution policies. Document No. 43 [2014] of the State Council first clarified the principle of no central government bailout. Document No. 88 [2016] of the General Office of the State Council proposed classified regulation but lacked detailed implementation rules. Document No. 50 [2017] of the Ministry of Finance promoted the construction of an explicit guarantee system through negative lists and market-oriented guarantee mechanisms. Document No. 35 and No. 47 issued in 2023 focused on targeted debt resolution and structural deleveraging, providing directional support to key provinces while strictly controlling new projects.

From the perspective of the guarantee analytical framework, the evolution of implicit guarantees exhibits phased characteristics. The debt restructuring clause in Document No. 43 was interpreted by the market as a continuation of implicit guarantees (Chen et al., 2020). Document No. 50 drove the shift of pricing fundamentals toward explicit guarantees (Ouyang & Wang, 2024). The targeted support in Document No. 35 still implies a logic of "selective support". This policy evolution directly affects the formation of urban investment bond spreads and reshapes market linkages by altering entity behaviors, providing a theoretical foundation for understanding the dynamic changes in network characteristics.

H_1: Although debt resolution policies such as Document 43, Document 50, and Document 35 aim to promote market-oriented clearing, varying degrees of implicit government guarantee still exist in their design.

2.2.2 Network Analysis:

Constructing a spatially correlated network of local government debt risks enables the accurate identification of key nodes and transmission channels of risk contagion from a structural perspective, revealing the dual characteristics of robustness and vulnerability coexisting in the system (Li, 2022). As a core tool for analyzing systemic correlation structures, complex network theory has formed a mature paradigm in research on the cross-entity and cross-market contagion mechanisms of financial risks. Numerous studies have confirmed that network topological features such as degree distribution and clustering coefficient are critical indicators for measuring risk accumulation and outbreak probability (Li et al., 2019; He et al., 2020). Therefore, examining local government debt risks within an integrated network framework is consistent with the evolutionary nature of debt risks and has solid methodological feasibility.

Against this background, this paper constructs two network indicators—issuer node degree and guarantor node degree—to capture how the central position of an issuer in the network structure affects its financing costs. Theoretically, this study introduces complex network analysis into the bond pricing framework, breaking through the limitations of traditional research that only focuses on individual characteristics, and reveals the important role of network externalities in risk pricing. Practically, accurately identifying network centrality features helps to warn against implicit risk contagion caused by joint guarantees, providing a decision-making basis for regulators to identify key nodes and establish a risk early warning system. On this basis, this paper proposes the following hypotheses:

H_2:The impact of network node degree on the issuance spread of urban investment bonds varies dynamically with the strength of implicit guarantees. Specifically, stronger implicit guarantees strengthen the negative effect of node degree on yield spreads, while weaker implicit guarantees amplify the positive effect of node degree on yield spreads. Moreover, explicit guarantees and implicit guarantees exhibit a substitution effect in this moderating mechanism.

III. RESEARCH DESIGN

3.1 Sample and Data Sources:

This paper collects issuance data of China's urban investment bonds, relevant enterprise data, and provincial local fiscal variables. The data are mainly obtained from the Wind Database, supplemented by the Resset Database. Short-term bonds with a maturity of less than 1 year and long-term bonds with a maturity of more than 10 years are excluded, and continuous variables are winsorized at the 1st and 99th percentiles. The final sample consists of 34,541 bonds issued from March 1, 2014

to March 31, 2024. Macroeconomic data are derived from the *China Statistical Yearbook*, the China Local Government Bond Information Disclosure Platform, and the China Money Network. The econometric software used is Stata 18.0.

3.2 Descriptive Statistical Analysis:

Based on the descriptive statistical results, the mean value of the explanatory variable issuance spread is 1.98 with a standard deviation of 1.11, indicating that the financing costs of urban investment bonds are somewhat differentiated but overall volatility is controllable.

In terms of network characteristics, the mean values of issuer node degree and guarantor node degree are 15.73 and 12.58 respectively, with maximum values reaching 322 and 473, and significantly high standard deviations. This reflects substantial differences in centrality among entities in the urban investment bond guarantee network, with a small number of core entities having far larger connection scales than ordinary entities.

In terms of bond micro characteristics, the mean values of total issuance amount and maturity are 8.35 billion yuan and 4.30 years, consistent with the mainstream issuance characteristics of urban investment bonds. The mean value of explicit guarantee is only 0.20, indicating that most urban investment bonds rely on implicit credit backing. In terms of macro and entity characteristics, the mean value of provincial fiscal self-sufficiency rate is 0.57, reflecting that most provinces rely on transfer payments for fiscal expenditure. The mean value of enterprise type is 0.99, confirming that nearly all urban investment issuers in the sample are state-owned enterprises. The mean registered capital is 8.002 billion yuan with large disparities, and the mean return on net assets is 1.02 with a negative minimum value, reflecting the large divergence in capital scale and weak overall profitability of urban investment platforms.

In addition, the mean values of SHIBOR and FR007 are stable with small volatility, indicating a relatively stable market liquidity environment during the sample period. Overall, the sample covers urban investment entities with different network positions and qualifications as well as diverse characteristics. The data distribution is consistent with the reality of the urban investment bond market, which can provide a reliable foundation for subsequent empirical analysis.

TABLE 2
DESCRIPTIVE STATISTICAL ANALYSIS

	Observations	Mean	Std	Min	Max
Issuance Spread	34541	1.98	1.11	0.32	5.1
Explicit Guarantee	34541	0.2	0.39	0	1
Issuer Node Degree	34541	15.73	20.84	0	322
Guarantor Node Degree	34541	12.58	48.67	0	473
Total Issuance Amount	34541	8.35	5.54	1	30
Maturity	34541	4.3	2.02	1	10
Fiscal Self-sufficiency Rate	34526	0.57	0.15	0.26	0.87
GDP	34532	38199.48	29510.36	2910.1	137008
PPI	34532	100.55	4.59	90.8	115.4
Fixed Asset Investment	33910	7.4	9.84	-27.4	43.9
Inflation Rate	34532	100.04	0.51	98.8	101.3
Growth Rate of Industrial Added Value	34532	6.78	4.86	-12.3	21.5
Money Supply	34541	10.24	1.64	8	13.7
Enterprise Type	34541	0.99	0.09	0	1
Registered Capital	34541	80.02	154.16	1	922.04
ROE	34541	1.02	3.07	-20.48	10.43
Net Asset Liability Ratio	34541	2.25	1.83	0.05	13.45
SHIBOR	34541	2.67	0.79	0.05	13.45
FR007	34541	2.34	0.5	1.4	4

3.3 Empirical Model Specification:

To examine the respective impacts of policy shocks and network characteristics on the issuance spread of urban investment bonds, this paper establishes the following three sets of empirical models.

3.3.1 Heterogeneity of Implicit Guarantees Implied by Policies:

To examine the changing role of explicit guarantees in the issuance pricing of urban investment bonds against the backdrop of evolving implicit guarantees, and to answer whether market-oriented credit enhancement mechanisms can better perform risk-sharing and reduce financing costs when government bailout expectations weaken, this paper first analyzes the dynamic changes in implicit guarantees. Various documents issued during the local government debt management reform provide a convenient setting for this study. We select Document 43, Document 88, Document 50, the outbreak of the COVID-19 pandemic, and Document 35 as time nodes, and use bonds issued six months before and after each event to conduct five sets of difference-in-differences regressions. The specific empirical model is as follows:

$$Spread_{i,p,t} = \alpha_0 + \alpha_1 Treated_i + \alpha_2 (Treated_i \times Post_t) + \sum_1^x \theta_x Control_{i,t} + \sum_1^x \mu_x Control_{i,p,t} + Time + Province + Industry + IssuerRating + \varepsilon_{i,t} \quad (1)$$

where *i* denotes the individual bond, and p denotes the province (autonomous region / municipality).The dependent variable $Spread_{i,p,t}$ is the difference between the coupon rate of bond i issued at time t and the yield of treasury bonds with the same maturity in the same period.The independent variable Treatedi is a dummy variable indicating whether the bond is an urban investment bond. To mitigate the endogeneity problem caused by omitted variables as much as possible, this paper includes a series of control variables $Control_{i,p,t}$ and $Control_{i,t}$, including bond-level and province-level control variables.The regression controls for monthly time fixed effects (Time), province fixed effects (Province), industry fixed effects (Industry), and issuer rating fixed effects (IssuerRating).

This paper expects that the strength of implicit guarantees for urban investment bonds will change with the issuance of relevant policies. Before the promulgation of policy documents, the market has high expectations of implicit guarantees for urban investment bonds, so their issuance spreads are significantly lower than those of the control group, corresponding to the coefficient α_1<0.As the implementation effects of different policies vary, if the degree of implicit guarantees weakens after the policy shock, the coefficient α_2 will be positive (α_2>0).Conversely, if the policy shock increases the degree of implicit guarantees, the coefficient α_2 will be negative (α_2<0).

3.3.2 The Substitution Relationship Between Explicit Guarantees and Implicit Guarantees:

From a theoretical perspective, the substitution relationship between explicit guarantees and implicit guarantees can be derived: when implicit guarantees exist, the importance and actual value of explicit guarantees are weakened, leading to differences in the impact of explicit guarantees on the issuance pricing of urban investment bonds and non-urban investment bonds. After understanding the changing trend of the degree of implicit guarantees, this paper constructs the following model to examine the changing role of explicit guarantees in the pricing of urban investment bonds:

$$Spread_{i,p,t} = \beta_0 + \beta_1 (CE_i \times Post_t) + \beta_2 (Treated_i \times Post_t) + \beta_3 CE_i + \beta_4 Treated_i + \sum_1^x \theta_x Control_{i,t} + \sum_1^x \mu_x Control_{i,p,t} + Time + Province + Industry + IssuerRating + \varepsilon_{i,t} \quad (2)$$

The core explanatory variable CE_i is a dummy variable for explicit guarantee, which takes the value of 1 if the bond is guaranteed, and 0 otherwise. $\beta_1$captures the impact of policy shocks on the effect of explicit guarantees. $\beta_2$reflects the change in credit spreads of urban investment bonds, i.e., the impact of policy shocks on implicit guarantee expectations.

Theoretically, as implicit guarantees are weakened by policies, explicit guarantees can better play a role in sharing bond risks and reducing financing costs. We therefore expect opposite signs for $\beta_1$and β_2: when implicit guarantee expectations are weakened, β_1<0 and β_2>0.

3.3.3 The Impact of Network Structure Information on Issuance Spread under the Effect of Implicit Guarantees:

To examine the basic effect of network characteristics on urban investment bonds and the moderating effect of policy periods on such impact, this paper constructs the following model:

$$Spread_{i,p,t} = \gamma_0 + \gamma_1\left(Treated_i \times \mathrm{I}_{\mathrm{Deegree}}\right) + \gamma_2\left(Treated_i \times \mathrm{G}_{\mathrm{Deegree}}\right) + \gamma_3 \mathrm{I}_{\mathrm{Deegree}} + \gamma_4 \mathrm{G}_{\mathrm{Deegree}} + \gamma_5 Treated_i + \sum_1^x \theta_x Control_{i,t} + \sum_1^x \mu_x Control_{i,p,t} + Time + Province + Industry + IssuerRating + \varepsilon_{i,t} \quad (3)$$

where coefficients $\gamma_1$and γ_2 are the coefficients of the core explanatory variables, identifying the pricing effects of issuer node degree and guarantor node degree on urban investment bonds; $\gamma_3$and $\gamma_4$represent the coefficients of issuer node degree and guarantor node degree on issuance spread across all bonds.

IV. REGRESSION RESULTS AND ANALYSIS

4.1 Analysis of the Heterogeneity of Implicit Guarantees Implied by Policies:

According to the regression results in Table 3, this paper analyzes the dynamic changes in implicit guarantees from the perspective of policy and pandemic shocks. Column (1) shows that the coefficient of urban investment bonds is significantly negative, which confirms the market expectation of government bailouts. The interaction terms in Columns (2) to (6) indicate that: After the issuance of Document 43, the issuance spread narrowed further, as the market interpreted "debt replacement" as a strengthening of implicit guarantees. Coupled with weaker-than-expected implementation and local economic stimulus, guarantee expectations were elevated. Although Document 88 attempted to regulate the debt boundary, its effect was limited. Implicit guarantees weakened significantly after the implementation of Document 50, which promoted the development of explicit guarantees by refining regulations, prohibiting commitment letters, and encouraging market-oriented guarantees. Expectations rebounded significantly after the pandemic shock, due to increased local financing dependence and slower policy implementation amid economic uncertainty. Document 35 reduced the issuance spread, as the market interpreted targeted restructuring and strict project controls as a demonstration of systemic risk management capacity, which in turn strengthened policy bailout expectations.

TABLE 3
HETEROGENEITY OF IMPLICIT GUARANTEES IMPLIED BY POLICIES

	(1)	(2)	(3)	(4)	(5)	(6)
	All Period	**Doc 43**	**Doc 88**	**Doc 50**	**the COVID-19**	**Doc 35**
Treated_Post	——	-0.420***	0.00307	0.180***	-0.237***	-0.392***
		(-4.90)	-0.04	-2.67	(-3.99)	(-13.90)
Treated	-0.200***	-0.055	-0.403***	-0.387***	-0.0875	0.0306
	(-17.25)	(-1.07)	(-5.17)	(-6.68)	(-1.63)	-1.05
Amount	-0.0218***	-0.0327***	-0.0107***	-0.0160***	-0.0182***	-0.0141***
	(-28.83)	(-10.52)	(-3.09)	(-6.01)	(-6.58)	(-9.49)
Maturity	0.0142*	0.0153	0.039	0.153***	0.101***	0.0484***
	-1.76	-0.41	-1.04	-5.23	-3.14	-2.58
RED	0.0844***	0.669***	0.403***	0.368***	0.329***	0.0602**
	-5.29	-5.14	-7.19	-8.97	-4.86	-2.41
PO	0.0662***	0.204*	0.343***	0.193***	0.103***	-0.0851***
	-6.9	-1.67	-8.57	-6.48	-3.36	(-4.19)
Fiscal Self-sufficiency Rate	0.847***	1.869	-0.0572	-6.224***	-0.985	3.203***
	-6.07	-1.35	(-0.05)	(-3.62)	(-1.14)	-3.58
Ln(GDP)	-0.0892***	0.544**	-0.274	0.31	-0.26	0.633**
	(-5.86)	-2.46	(-0.65)	-1.34	(-0.99)	-2.18
PPI	0.00923***	0.000698	-0.0182	0.0118	-0.00545	-0.0265**
	-4.67	-0.03	(-1.41)	-0.85	(-0.29)	(-2.10)
Fixed Asset Investment	-0.00487***	0.00114	0.000697	0.00347	-0.0047	-0.00532**
	(-7.43)	-0.12	-0.1	-0.6	(-1.47)	(-2.07)

Inflation Rate	0.00694	0.0749	-0.0247	0.151**	0.143**	-0.0625
	-0.43	-1.06	(-0.25)	-2.3	-2.5	(-1.04)
Growth Rate of Industrial Added Value	0.00171*	-0.00531	0.014	0.0113	0.00965**	-0.00493***
	-1.69	(-0.40)	-1.19	-1.45	-2.15	(-2.61)
Money Supply	0.126***	0	0	0.453**	0	0
	-14.17	0	0	-2.53	0	0
Enterprise Type	-0.157**	0	-0.651**	0.82	-1.493***	-0.314***
	(-2.19)	0	(-2.35)	-1.6	(-5.23)	(-3.53)
Registered Capital	-0.000357***	0.000112	-0.000319**	-0.000369***	-0.000385***	-0.000436***
	(-13.25)	-0.67	(-2.33)	(-3.05)	(-4.71)	(-8.85)
ROE	-0.0215***	-0.00693	-0.00555	-0.0150**	-0.0225***	-0.0225***
	(-15.14)	(-1.03)	(-0.61)	(-2.20)	(-3.38)	(-10.33)
Net Asset Liability Ratio	-0.0278***	-0.00338	-0.0200*	-0.0299***	-0.0380***	-0.0280***
	(-11.75)	(-0.27)	(-1.73)	(-2.82)	(-3.61)	(-7.64)
Shibor	-0.00662	-0.0105	1.003**	0.900***	0.037	-0.15
	(-0.25)	(-0.06)	-2.21	-4.58	-0.29	(-0.61)
FR007	0.127***	0.0667*	0.0463	0.0785*	0.112	0.0991***
	-8.12	-1.67	-0.7	-1.66	-1.62	-3.74
Constant	-0.275	-10.53	6.502	-22.67***	-6.983	2.778
	(-0.16)	(-1.27)	-0.67	(-2.98)	(-1.05)	-0.43
Time	YES	YES	YES	YES	YES	YES
Province	YES	YES	YES	YES	YES	YES
Industry	YES	YES	YES	YES	YES	YES
IssuerRating	YES	YES	YES	YES	YES	YES
N	31312	973	1091	1407	3149	8061
R^2	0.61	0.657	0.579	0.505	0.601	0.579

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Robust standard errors are reported in parentheses. Columns (2) to (6) correspond to newly issued bonds before and after the policy shocks. The time window restriction results in a decrease in sample size, the same for the following tables.

4.2 The Substitution Relationship between Explicit Guarantees and Implicit Guarantees:

According to Table 4, the phased regression reveals the dynamic interaction between explicit guarantees and implicit guarantees under different policy shocks. In the full-sample regression, the main effect of implicit guarantees is significantly negative, while that of explicit guarantees is significantly positive, confirming the dominant role of implicit guarantees and their substitution relationship with explicit guarantees in the long-run equilibrium. However, this framework undergoes structural changes under policy shocks.

Specifically: During the period of Document 43, the interaction term of implicit guarantees is significantly negative, while that of explicit guarantees is positive but insignificant, indicating that market expectations of implicit guarantees strengthened instead, and explicit guarantees did not play an incremental role. During the period of Document 88, the interaction terms of both types of guarantees are insignificant, as the market adopted a wait-and-see attitude toward the policy. During the period of Document 50, the interaction term of implicit guarantees is significantly positive (expectations weakened), while that of explicit guarantees is negative, suggesting that explicit guarantees began to show a potential substitution effect after the weakening of implicit guarantees, although a market consensus had not yet been formed. During the COVID-19 pandemic and the period of Document 35, the interaction terms of both guarantees are significantly negative, indicating that in the face of systemic risks or overall debt restructuring policies, investors no longer distinguish between the two, but regard them as superimposed risk mitigation tools, resulting in a synergistic enhancement effect.

TABLE 4
THE SUBSTITUTION RELATIONSHIP BETWEEN EXPLICIT GUARANTEES AND IMPLICIT GUARANTEES

	(1)	(2)	(3)	(4)	(5)	(6)
	All Period	Doc 43	Doc 88	Doc 50	the COVID-19	Doc 35
Treated_Post	——	-0.416***	-0.0171	0.129*	-0.209***	-0.310***
		(-4.74)	(-0.21)	-1.92	(-3.51)	(-10.98)
CE_Post	——	0.0369	0.09	-0.0415	-0.125	-0.436***
		-0.27	-1.09	(-0.61)	(-1.59)	(-10.07)
Treated	-0.204***	-0.0311	-0.368***	-0.354***	-0.118**	-0.00992
	(-17.66)	(-0.61)	(-4.69)	(-6.25)	(-2.19)	(-0.34)
CE	0.0689***	-0.307***	-0.465***	-0.430***	0.249***	0.264***
	-4.96	(-5.61)	(-6.97)	(-7.23)	-3.86	-7.16
Control	YES	YES	YES	YES	YES	YES
N	31312	973	1091	1407	3149	8061
R^2	0.611	0.671	0.607	0.557	0.604	0.586

4.3 The Impact of Network Structure Information on Issuance Spread under Implicit Guarantees:

Table 5 presents the dynamic impacts of issuer node degree and guarantor node degree on bond issuance spreads under implicit guarantees. The full-sample results show that the effects of node degree differ significantly between urban investment bonds and non-urban investment bonds: for non-urban investment bonds, the effect of guarantor node degree is insignificant, whereas for urban investment bonds, it significantly reduces spreads, reflecting the credit endorsement effect under implicit guarantees.

In terms of sub-periods: During the Document 43 period, the market interpreted debt replacement as a strengthening of implicit guarantees. The spread-increasing effect of issuer node degree in non-urban investment bonds was offset in urban investment bonds, and the spread-suppressing effect of guarantor node degree was weakened. During the Document 88 period, although the policy attempted to clarify the debt boundary, implicit guarantees still persisted, and the market-based credit role of guarantor node degree began to emerge. During the Document 50 period, the prohibition of implicit guarantees promoted market-oriented transformation. The effect of node degree returned to market logic, the difference in the impact of issuer node degree narrowed, and the suppressing effect of guarantor node degree became more significant for urban investment bonds. During the pandemic shock period, expectations of implicit guarantees rebounded. The coefficients of both issuer node degree and guarantor node degree tended to be negative, but the latter was insignificant, confirming that stronger implicit guarantees suppress the signal role of node degree. During the Document 35 period, which focused on debt resolution, expectations of implicit guarantees rebounded slightly, and the effect of node degree became divergent again.

Overall, the impact of node degree is highly linked to the strength of implicit guarantees: the stronger the implicit guarantee, the more the market-based credit signal of node degree is suppressed; the weaker the implicit guarantee, the more significant its pricing effect. This confirms the transition of credit pricing for urban investment bonds from “implicit guarantee dependence” to “market-oriented credit dominance”.

TABLE 5
THE IMPACT OF NETWORK STRUCTURE INFORMATION ON ISSUANCE SPREAD UNDER IMPLICIT GUARANTEES

	All Period	Doc 43	Doc 88	Doc 50	the COVID-19	Doc 35
I_Degree	0.00105***	0.0196**	0.00889*	0.00355	0.00540***	-0.00114**
	-4.26	-2.47	-1.65	-0.84	-3.22	(-2.55)
Treated_I_Degree	0.00107***	-0.00471	-0.000806	-0.0002	-0.00562***	0.00195***
	-3.29	(-0.43)	(-0.12)	(-0.04)	(-2.76)	-3.17
G_Degree	0.000163	-0.00540***	-0.0221***	-0.0241**	-0.000232	0.00121***
	-0.6	(-4.57)	(-3.49)	(-2.21)	(-0.28)	-2.63
Treated_G_Degree	-0.00111***	0.00126	0.0150**	0.0184*	-0.000429	-0.00181***
	(-3.92)	-1.01	-2.34	-1.69	(-0.45)	(-3.74)
Treated	-0.219***	-0.11	-0.384***	-0.275***	-0.169***	-0.154***
	(-16.25)	(-1.43)	(-4.36)	(-4.27)	(-3.39)	(-5.43)
Control	YES	YES	YES	YES	YES	YES
N	31312	973	1091	1407	3149	8061
R^2	0.612	0.663	0.601	0.553	0.6	0.572

V. ROBUSTNESS CHECKS

5.1 The Substitution Relationship between Network Structure, Explicit Guarantees and Implicit Guarantees Acknowledgement:

Table 6 examines the substitution relationship between explicit guarantees and implicit guarantees under network structure. In the left column without explicit guarantees, the coefficient of Treated is significantly negative, confirming that implicit guarantees reduce the issuance spreads of urban investment bonds. The main effect of issuer node degree is significantly positive, and the interaction term with urban investment bonds is also significantly positive, indicating that its risk signal is still partially transmitted. The main effect of guarantor node degree is insignificant, whereas its interaction term with urban investment bonds is significantly negative, reflecting the strengthened credit endorsement of the guarantor's network status under implicit guarantees. In the right column with explicit guarantees included, the coefficient of CE is significantly positive, reflecting that explicit guarantees are associated with higher spreads among non-urban investment bonds. The interaction term of issuer node degree becomes significantly negative, while the interaction term of guarantor node degree becomes insignificant. This suggests that implicit guarantees mask the risk signals of both explicit guarantees and node degree, and their network value is dominated by the combined effect.

Overall, in the presence of implicit guarantees, the effect of explicit guarantees is weakened, and the credit signal of node degree is also diminished. This is consistent with the theoretical expectation that the stronger the implicit guarantee, the more limited the independent role of explicit guarantees.

TABLE 6
THE SUBSTITUTION RELATIONSHIP BETWEEN NETWORK STRUCTURE, EXPLICIT GUARANTEES AND IMPLICIT GUARANTEES

	Spread		Spread
Treated	-0.219***	CE	0.159***
	(-15.78)		-9.53
I_Degree	0.00169***	I_Degree	0.00169***
	-6.01		-6.01
G_Degree	-0.0158	G_Degree	-0.0158
	(-1.34)		(-1.34)
Treated _G_Degree	-0.00141***	Treated _G_Degree	0.0159
	(-3.76)		-1.34
Treated _I_Degree	0.000717**	Treated _I_Degree	-0.00150***
	-2.05		(-3.07)
Constant	-0.632	Constant	-0.632
	(-0.36)		(-0.36)
Control	YES	Control	YES
N	31312	N	31312
R^2	0.614	R^2	0.614

5.2 Parallel Trends Test:

The results of the parallel trends test show that before the policy shocks of Document 43, Document 50, the pandemic, and Document 35, the coefficients of the interaction terms between the treatment variable and the monthly dummy variables are all insignificant. The treatment group and the control group exhibit no systematic difference in the credit spreads of urban investment bonds, which satisfies the parallel trends assumption.

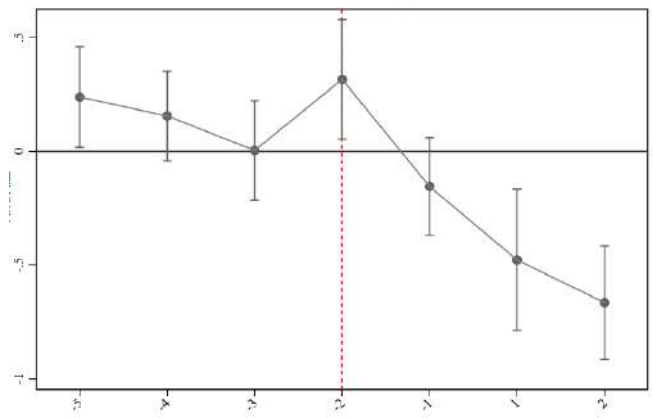


FIGURE 2 (1) Parallel Trends Test for Document 43

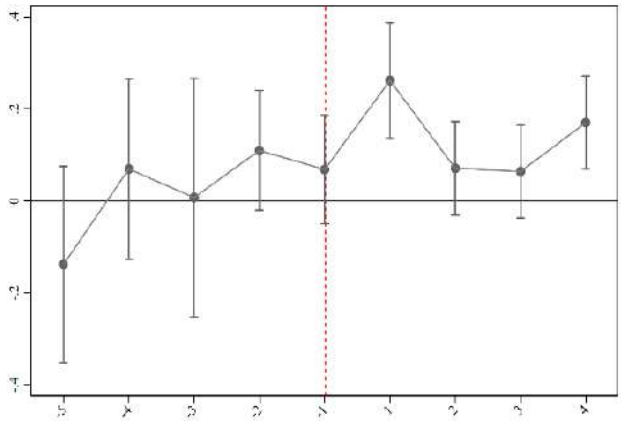


FIGURE 2 (2) Parallel Trends Test for Document 50

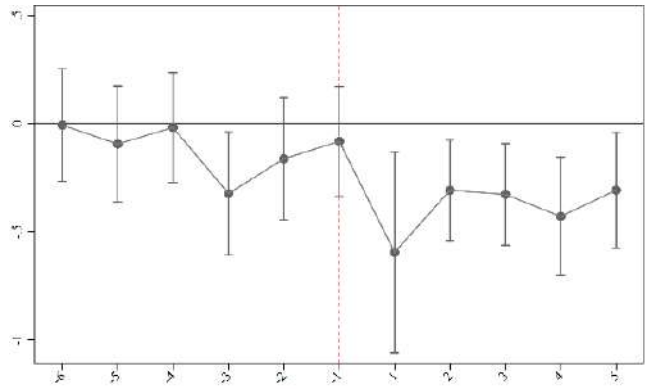


FIGURE 2 (3) Parallel Trends Test for the COVID-19

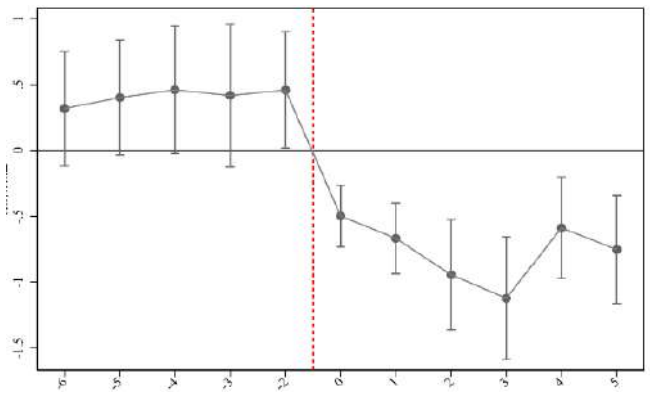


FIGURE 2 (4) Parallel Trends Test for Document 35

5.3 PSM-DID Test:

To mitigate the endogeneity problem caused by sample self-selection, this paper further employs the PSM-DID method for re-estimation. We take the total issuance amount, maturity, puttable provision, and callable provision of bonds as covariates. The propensity score is estimated using a Logit model, and 1:1 and 1:3 nearest-neighbor matching with replacement is implemented. Benchmark regression analysis is then conducted using the matched sample. The results show that the sign and significance level of the coefficients of the core explanatory variables are generally consistent with those of the benchmark regression, indicating that the estimated policy effects are not significantly affected by sample selection bias. The findings support the original conclusions.

VI. CONCLUSIONS

This paper constructs a dynamic network indicator system to systematically examine the interactive effects of policies (Document 43, Document 50, pandemic shock, and Document 35) and network structure on the issuance spreads of urban investment bonds. The conclusions are as follows:

First, the policy effects exhibit significant heterogeneity, depending on the market's interpretation of policy connotations. Debt replacement under Document 43 was regarded as credit endorsement, thus reducing spreads; Document 50 effectively transmitted the signal of breaking rigid repayment; and targeted support under Document 35 strengthened expectations of government bailouts. Second, the pricing weight of network characteristics is structurally reshaped with the policy environment: issuer node degree shifts from a policy transmission channel to a market-oriented signal, while guarantor node degree reverses from credit enhancement to a risk transmission hub, confirming the transition of the pricing benchmark from relationship dependence to risk-based pricing. Third, the pricing of urban investment bonds systematically deviates from traditional financial theories. Variables related to government credit, such as enterprise ownership and issuer rating, become central to pricing, highlighting the dominant role of "government credit discount".

Accordingly, we propose the following recommendations: improve the policy transmission mechanism and stabilize market expectations through regular communication; establish a network risk monitoring system, incorporate node degree into the regulatory framework, and promote the diversification of the guarantee system; deepen the reform of information disclosure and credit rating, strengthen the evaluation of individual solvency, and provide institutional support for market-oriented pricing. This study provides a new perspective for understanding the pricing mechanism of urban investment bonds and also offers empirical evidence for debt risk prevention and control.

ACKNOWLEDGEMENT

The authors have no acknowledgements to declare.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] Qiu, Z. G., Wang, Z. Y., & Wang, Z. (2022). Local government debt replacement and new implicit debt: Analysis based on the issuance scale and pricing of urban investment bonds. *China Industrial Economics*, (4), 42–60.

[2] Cao, J. (2023). Divergence in beliefs in rigid repayment: Implicit guarantee expectations and urban investment bond pricing. *The Journal of World Economy, 46*(6), 85–107.

[3] Altman, E. I. (1990). *Corporate financial distress*. Wiley. pp. 212–237.

[4] Collin-Dufresne, P., Goldstein, R. S., & Martin, J. S. (2001). The determinants of credit spread changes. *Journal of Finance, 56*(6), 2177–2207.

[5] Pan, J., Wang, L. L., & Shen, X. F. (2015). Financial ecological environment and local government debt financing cost — Empirical test based on provincial urban investment bond data. *Accounting Research*, (6), 34–41, 96.

[6] Yan, S. H., & Ma, W. J. (2024). Fiscal sustainability and the resolution of local government debt risks: A perspective based on special bond spreads. *Inquiry into Economic Issues*, (11), 1–20.

[7] Fons, J. S. (1994). Using default rates to model the term structure of credit risk. *Financial Analysts Journal, 50*(5), 25–32.

[8] Kalotay, A. J. (1997). Optimum high yield bond calling. *Extra Credit, 8*(1), 34–39.

[9] Zhang, X. Y., & Jiao, J. (2017). Research on the impact of guarantees on bond issuance spreads. *Collected Essays on Finance and Economics*, (2), 48–57.

[10] Bo, L., Yao, H., Mear, F. C. J., & [Additional authors if available]. (2023). How does China's new budget law affect the pricing of local government bonds? *Applied Economics Letters, 30*(7), 937–943.

[11] Chen, F., Liu, Q., & Wang, Y. (2020). Why do firms issue guaranteed bonds? *Journal of Banking & Finance, 119*, Article 105396.

[12] Ouyang, Y. F., & Wang, Q. S. (2024). Urban investment bond pricing based on a comparison between implicit and explicit guarantees. *Journal of Financial Research*, (8), 77–94.

[13] Liang, W. J., Ji, X. G., Ning, B., et al. (2025). The credit aid effect of counterpart assistance — Evidence from municipal investment bond issuance pricing. *Economic Quarterly, 25*(3), 734–749.

[14] Tong, Y., Li, X., & Xu, G. Q. (2024). Shareholding in listed firms by local government financing vehicles and pricing of urban investment bonds. *Nankai Economic Studies*, (4), 107–126.

[15] Mao, J., Han, R. X., & Liu, C. (2024). A new mechanism of debt growth in local government financing vehicles: From the perspective of guarantee networks. *Economic Research Journal, 59*(1), 72–92.

[16] Wu, D. S., Cao, Y., Tang, C., et al. (2021). Research on debt risk and risk contagion network under classified control. *Journal of Management World, 37*(4), 35–54.

[17] Wang, Y. Q., Chen, Y. H., & Du, J. L. (2016). Soft budget constraint and local government debt default risk in China: Evidence from financial markets. *Economic Research Journal, 51*(11), 96–109.

[18] Luo, R. H., & Liu, J. J. (2016). Is implicit guarantee by local governments really effective? — A test based on urban investment bond issuance pricing. *Journal of Financial Research*, (4), 83–98.

[19] Li, T. (2022). *Research on network contagion and spatial effects of local government debt risk* [Doctoral dissertation, Southwestern University of Finance and Economics].

[20] He, F., Teng, X. X., & Wang, S. W. (2020). Research on the complex network structure and systemic risk characteristics of local government bonds. *Statistics & Decision, 36*(4), 136–140.

[21] Li, S. W., Wen, S. H., Wang, L., et al. (2020). Research on the evolution characteristics of financial institutions' relevance from a multilayer network perspective. *Chinese Journal of Management Science, 28*(12), 35–43.

International Journal of Engineering Research & Science (IJOER)
ISSN: [2395-6992]
[Vol-12, Issue-3, March- 2026]
Website: www.ijoer.com
Journal DOI: 10.25125/engineering-journal

Review of Processing and Manufacturing Challenges in the Fabrication of Ceramic Matrix Composites

Dr. Balasubramanyam. N[1*]; D. Jyosthna[2]; A.V.N.S. Kiran[3]

Dept. of Mechanical Engineering, Sri Venkateswara University College of Engineering, Tirupati, Andhra Pradesh, India-517502
*Corresponding Author

Received: 08 March 2026/ Revised: 17 March 2026/ Accepted: 23 March 2026/ Published: 31-03-2026

Abstract— *Ceramic Matrix Composites (CMCs) have emerged as essential advanced materials for high-temperature and high-performance applications, offering superior thermal stability, lightweight design, oxidation protection, and enhanced fracture resistance compared to traditional monolithic ceramics. Their growing use in demanding applications, such as aerospace propulsion systems, automotive components, and advanced energy technologies, is driven by their unique ability to meet these stringent requirements. However, significant technical and economic obstacles continue to limit their widespread industrial implementation. This paper presents a critical review of the key barriers to CMC development, including inherent pseudo-ductility limits, complexities in fabrication processes, high production costs, challenges in fibre–matrix interface engineering, susceptibility to environmental degradation, and the lack of standardized design methodologies and material databases. The study analyses recent advancements in processing technologies, interfacial design, and environmental protection strategies that aim to improve CMC performance, reliability, and manufacturability. By establishing the current limitations of CMCs, this work identifies future research opportunities necessary to accelerate their adoption in next-generation engineering systems.*

Keywords— Ceramic Matrix Composites, Chemical Vapour Infiltration, Fibre-Matrix Interphase, Environmental Barrier Coatings, Processing Challenges.

I. INTRODUCTION

Ceramic Matrix Composites (CMCs) are advanced engineered materials that combine high-strength ceramic fibres with a ceramic matrix to achieve high-temperature performance while mitigating the inherent brittleness and low fracture toughness of monolithic ceramics. Traditional ceramics, despite possessing high hardness, oxidation resistance, and thermal stability, are highly susceptible to catastrophic failure due to their inability to resist crack propagation. The introduction of reinforcing fibres with engineered interfaces enables CMCs to achieve toughening through crack deflection, fibre bridging, and fibre pull-out, resulting in enhanced damage tolerance and reliability (Evans & Marshall, 1989).

In recent decades, CMCs have gained considerable attention in high-performance engineering applications, particularly in aerospace, defence, and energy sectors. Their capability to function at temperatures above 1000°C in hostile oxidizing environments enables their use in gas turbine engine components, combustor liners, exhaust nozzles, and spacecraft thermal protection systems. CMCs provide three key benefits over traditional metallic alloys: they enhance propulsion system efficiency and decrease emissions through their lighter weight, ability to withstand higher temperatures, and reduced need for cooling (Naslain, 2004). These attributes align with current environmental demands for technologies that consume less energy and produce fewer harmful emissions.

The performance of CMCs is governed by their microstructural design, which dictates material properties. A critical design element is the weak fibre–matrix interphase, which enables controlled de-bonding during crack propagation, thereby absorbing energy. This interphase design allows cracks to follow the fibre–matrix interface rather than propagating directly through the

fibres, preventing unexpected catastrophic failure. Consequently, CMCs display pseudo-ductile behaviour—a major enhancement over the brittle characteristics of monolithic ceramics. However, achieving the optimal balance between interfacial bonding strength and de-bonding ability remains an ongoing focus of materials engineering.

Despite these advantages, several critical factors limit the widespread adoption of CMCs. The fabrication processes, such as Chemical Vapour Infiltration (CVI), Polymer Infiltration and Pyrolysis (PIP), and Melt Infiltration (MI), are time-consuming and expensive, requiring precise control to produce consistent microstructures. The inherent thermal expansion coefficient mismatch between fibre and matrix phases introduces residual stresses that can lead to micro-cracking during manufacturing and service (Evans, 1990). Furthermore, environmental degradation due to oxidation and moisture attack at high temperatures can compromise material integrity and reduce operational life. Beyond materials-specific challenges, industrial adoption is obstructed by the absence of standardized design methodologies, a lack of comprehensive long-term performance data, and difficulties in scaling manufacturing processes. Addressing these obstacles requires sustained research into advanced material designs, improved processing methods, and cost-effective manufacturing approaches.

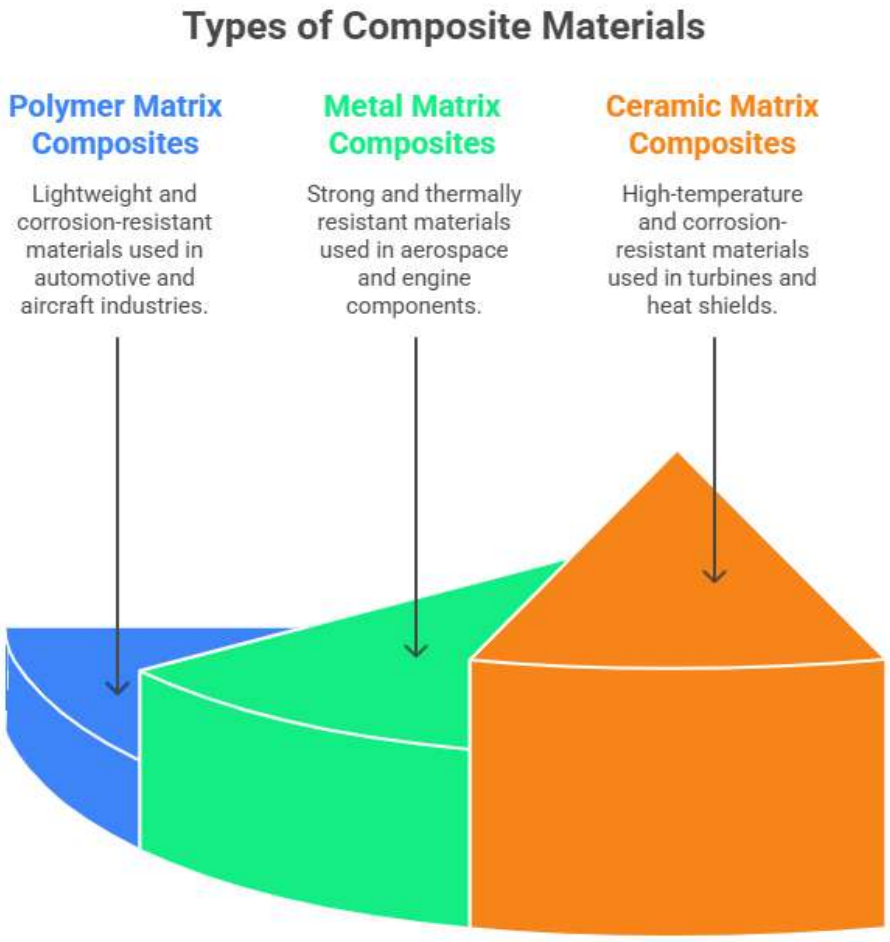


FIGURE 1: Types of Composite Materials

II. CHALLENGES IN CMC DEVELOPMENT

The development of CMCs is fundamentally challenged by their inherent mechanical behavior. While fibre reinforcement significantly improves toughness compared to monolithic ceramics, CMCs can still experience failure under tensile or impact loading due to microstructural defects that serve as crack initiation sites. Unlike metals, which exhibit plastic deformation, CMCs rely on complex mechanisms such as fibre pull-out and crack bridging to dissipate energy. Controlling these mechanisms to prevent unstable crack propagation requires precise engineering of the fibre–matrix interface (Marshall & Cox, 1988).

2.1 Processing Complexity:

A major obstacle lies in the manufacturing methods themselves. The primary fabrication routes—CVI, PIP, and MI—each present distinct challenges. CVI, while capable of producing high-quality matrices, is characterized by long processing times (often hundreds of hours) and difficulty in achieving uniform densification in thick sections due to diffusion-limited transport. PIP offers lower processing temperatures but suffers from high porosity and multiple infiltration cycles, leading to process variability. MI enables rapid densification but involves high-temperature processing that can degrade fibre properties and introduce residual thermal stresses due to molten metal infiltration (Krenkel, 2008). Achieving uniform matrix densification while controlling porosity and eliminating defects remains a persistent challenge, often resulting in material property variations that complicate certification for safety-critical applications.

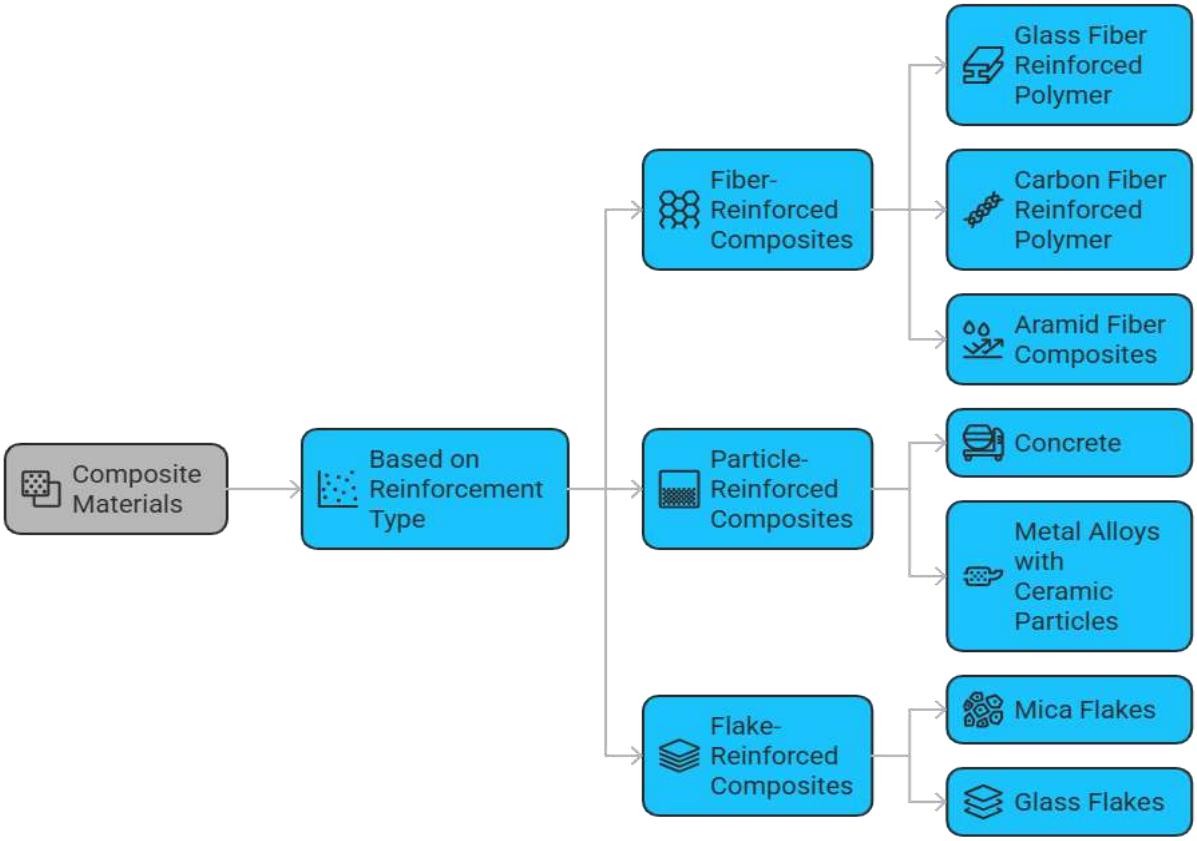


FIGURE 2: Types of Composites based on Reinforcement

2.2 High Production Costs:

The high production costs of CMCs represent a significant barrier to their widespread adoption across industries beyond aerospace and defence. The cost is driven primarily by the high price of continuous ceramic fibres (e.g., SiC-based fibres such as Hi-Nicalon or Tyranno), which require energy-intensive processing routes and yield low production volumes. Additionally, the slow, multi-step nature of densification processes contributes to elevated manufacturing costs. Consequently, the economic viability of CMCs is currently restricted to applications where performance justifies the premium, such as turbine blades and combustor liners (Bansal & Lamon, 2014). Ongoing research focuses on reducing costs through faster processing cycles, alternative fibre architectures, and the development of more affordable precursor systems.

2.3 Fibre–Matrix Interface Engineering:

The fibre–matrix interface is critical to mechanical performance, yet its optimization remains challenging. The interface must be sufficiently strong to transfer loads from the matrix to the fibres but weak enough to promote controlled de-bonding and fibre pull-out, which are essential for toughness. This balance is difficult to achieve because high-temperature processing steps can alter interface properties, promote chemical reactions between fibres and matrix, or degrade fibre strength. The development of stable interphase materials, such as boron nitride (BN) or multilayered coatings, is an active area of research aimed at maintaining this delicate balance under processing and service conditions.

2.4 Environmental Degradation:

Environmental degradation further complicates CMC development. In high-temperature, oxidizing environments, particularly those containing water vapour, the silica-based phases commonly present in CMCs undergo volatilization. This can lead to accelerated oxidation of the fibre–matrix interphase, resulting in mechanical property loss and reduced service life (Opila et al., 2002). While Environmental Barrier Coatings (EBCs) have been developed to mitigate these effects, ensuring long-term coating stability and adhesion under thermal cycling and high-velocity gas flow remains an unresolved issue.

2.5 Thermal Residual Stresses and Standardization Gaps:

The mismatch in coefficient of thermal expansion (CTE) between fibres and matrix generates residual stresses during cooling from processing temperatures. These stresses can initiate micro-cracks, which serve as preferential pathways for oxidation and reduce composite strength. Addressing this issue requires careful selection of material systems and process optimization to minimize residual stress accumulation (Evans, 1990). Finally, the absence of standardized material databases and design methodologies prevents engineers from confidently implementing CMCs in load-bearing structures. The lack of widely accepted testing protocols, long-term creep data, and fatigue life models creates uncertainty in design and certification processes (Bansal & Lamon, 2014).

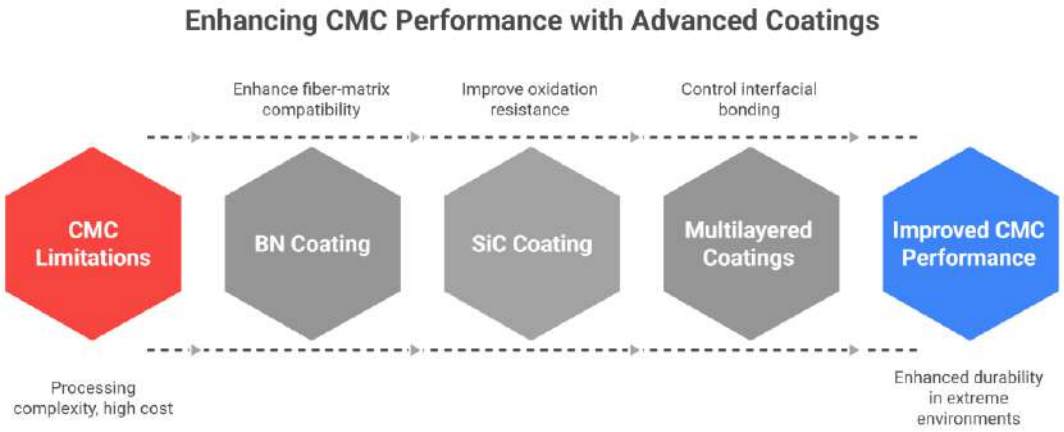


FIGURE 3: Enhancing CMC Performance with Advanced Coatings

III. CURRENT RESEARCH DIRECTIONS

Current research on CMCs is focused on overcoming the processing, cost, and environmental durability challenges outlined above. One of the most active areas is the development of advanced fiber coatings and interphases. Researchers are investigating materials such as boron nitride (BN), silicon carbide (SiC), and nano-engineered multilayered coatings to create interfaces that provide oxidation resistance while retaining the necessary mechanical properties for crack deflection and energy dissipation (Naslain, 2004).

3.1 Hybrid and Advanced Processing:

To address processing limitations, hybrid fabrication techniques are being explored. For example, combining CVI with MI aims to leverage the uniform coating capability of CVI with the rapid densification of MI, resulting in composites with improved mechanical properties and reduced manufacturing defects. Additionally, the integration of process monitoring and automated control systems is being pursued to enhance repeatability and reduce variability in production (Krenkel, 2008).

3.2 Additive Manufacturing:

Additive manufacturing (AM) has emerged as a potential paradigm shift in CMC fabrication. Techniques such as direct ink writing, binder jetting, and stereolithography are being studied for their ability to produce complex, near-net-shape geometries that are difficult or impossible to achieve with conventional processing. AM offers the potential to reduce material waste, shorten production lead times, and enable novel architectures with optimized fibre placement. However, significant challenges remain, including achieving high fibre volume fractions, maintaining fibre alignment, and obtaining full densification with minimal porosity after post-processing.

3.3 Environmental Barrier Coatings (EBCs):

The durability of CMCs in combustion environments is being addressed through the development of next-generation EBCs. Current research is focused on rare-earth silicate-based multilayer coatings designed to provide long-term protection against oxidation, corrosion, and water-vapour-induced recession (Opila et al., 2002). These advanced coating systems are critical for ensuring the operational lifespan of CMC components in aerospace engines, where they must withstand extreme thermal gradients and high-velocity gas environments.

3.4 Computational Modeling and Data-Driven Design:

Computational modelling now plays a central role in accelerating CMC development. Multi-scale modelling approaches are being developed to link processing conditions to microstructural evolution and ultimately to macroscopic mechanical properties. These models enable the prediction of damage initiation, thermal stress distribution, and long-term performance under complex loading scenarios. Furthermore, the integration of artificial intelligence and machine learning is enabling data-driven materials design, allowing researchers to explore vast processing–structure–property spaces more efficiently than traditional trial-and-error approaches.

3.5 Cost Reduction and Sustainability:

A major focus of ongoing research is the reduction of total manufacturing costs. Strategies include the development of faster densification cycles, the use of less expensive fibre architectures such as woven fabrics instead of unidirectional tapes, and the exploration of recyclable or more sustainable precursor materials. These efforts aim to make CMCs economically viable for a broader range of applications, including automotive and industrial gas turbines.

Finally, collaborative initiatives between academia, industry, and government agencies are essential for translating laboratory-scale innovations to industrial practice. These partnerships support the development of standardized databases, shared testing methodologies, and qualification frameworks necessary for widespread adoption. Collectively, current research directions aim to deliver CMC materials with improved performance, higher reliability, and lower production costs, paving the way for their integration into future engineering systems.

IV. CONCLUSION

Ceramic Matrix Composites offer exceptional properties that make them highly suitable for demanding high-temperature and high-performance applications. However, their widespread adoption is constrained by a combination of technical and economic challenges, including inherent limitations in pseudo-ductile behavior, complex and costly processing routes, difficulties in interface engineering, susceptibility to environmental degradation, and the absence of standardized design methodologies. Addressing these challenges requires continued advances in processing technologies, such as hybrid fabrication and additive manufacturing, alongside innovations in interface design and environmental barrier coatings. The integration of computational modeling and data-driven approaches will be critical to accelerating the optimization of processing–structure–property relationships. Through sustained multidisciplinary research and closer collaboration between industry and academia, the barriers to adoption can be systematically overcome, enabling CMCs to fulfil their potential as enabling materials for next-generation engineering systems

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Bansal, N. P., & Lamon, J. (2014). *Ceramic matrix composites: Materials, modeling and technology*. Wiley.

[2] Evans, A. G., & Marshall, D. B. (1989). The mechanical behavior of ceramic matrix composites. *Acta Metallurgica, 37*(10), 2567–2583.

[3] Evans, A. G. (1990). Perspective on the development of high-toughness ceramics. *Journal of the American Ceramic Society, 73*(2), 187–206.

[4] Krenkel, W. (2008). *Ceramic matrix composites: Fiber reinforced ceramics and their applications*. Wiley-VCH.

[5] Marshall, D. B., & Cox, B. N. (1988). A J-integral method for calculating steady-state matrix cracking stresses in composites. *Mechanics of Materials, 7*(2), 127–133.

[6] Naslain, R. (2004). Design, preparation and properties of non-oxide CMCs for application in engines and nuclear reactors. *Composites Science and Technology, 64*(2), 155–170.

[7] Opila, E. J., Smialek, J. L., Robinson, R. C., Fox, D. S., & Jacobson, N. S. (2002). Oxidation and volatilization of silica-forming ceramics in water vapor. *Journal of the American Ceramic Society, 85*(3), 659–665.

[8] Curtin, W. A. (1991). Theory of mechanical properties of ceramic-matrix composites. *Journal of the American Ceramic Society, 74*(11), 2837–2845.

[9] Zok, F. W. (2001). Developments in oxide fiber composites. *Journal of the American Ceramic Society, 84*(5), 933–944.

[10] Chawla, K. K. (2012). *Composite materials: Science and engineering* (3rd ed.). Springer.

[11] Schneider, H., & Komarneni, S. (2005). *Ceramic matrix composites: Materials, modeling and technology*. Wiley.

[12] Naslain, R. (2005). SiC-matrix composites: Nonbrittle ceramics for thermostructural application. *International Journal of Applied Ceramic Technology, 2*(2), 75–84.

[13] Singh, M., Levine, S. R., & Lowden, R. A. (2003). Processing, structure, and properties of silicon carbide fiber-reinforced ceramic matrix composites. *Journal of the American Ceramic Society, 86*(8), 1341–1356.

[14] Levi, C. G. (2004). Emerging materials and processes for thermal barrier systems. *Current Opinion in Solid State and Materials Science, 8*(1), 77–91.

[15] DiCarlo, J. A., Bhatt, R. T., & Roode, M. (2004). Continuous fiber ceramic composite materials for aerospace applications. *Journal of the American Ceramic Society, 87*(3), 321–330.

[16] Fitzer, E., & Gadow, R. (1986). Fiber-reinforced silicon carbide. *American Ceramic Society Bulletin, 65*(2), 326–335.

[17] Padture, N. P., Gell, M., & Jordan, E. H. (2002). Thermal barrier coatings for gas-turbine engine applications. *Science, 296*(5566), 280–284.

International Journal of Engineering Research & Science (IJOER)
ISSN: [2395-6992]
[Vol-12, Issue-3, March- 2026]
Website: www.ijoer.com
Journal DOI: 10.25125/engineering-journal

Research on Interpretable Loan Approval Identification Using Multi-Dimensional Features

Tiantian Lu

Economic and Social Statistics, Guizhou University of Finance and Economics, Guiyang-550025, China

Received: 09 March 2026/ Revised: 15 March 2026/ Accepted: 24 March 2026/ Published: 31-03-2026

Abstract— *Loan approval is a pivotal component of financial risk control. Current research often relies on black-box prediction models and lacks in-depth exploration of multidimensional features and interpretability, resulting in shortcomings in model interpretability and robustness. To address this, this paper proposes an Interpretable Loan Approval Identification Model based on Multidimensional Features (ILA-MDF). The ILA-MDF model is constructed using the CatBoost algorithm and is compared with benchmark models such as Random Forest in experimental evaluations. The results indicate that the proposed model performs optimally across six metrics—Accuracy, Precision, Recall, F1-score, AUC, and MCC—achieving values of 93.33%, 87.19%, 82.03%, 84.53%, 97.77%, and 80.34%, respectively. Furthermore, the SHAP framework is introduced to analyze the key factors influencing model decisions. Interpretability analysis reveals that the probability of loan approval increases significantly when the loan interest rate exceeds 14% or when the previous loan default indicator is 0. The ILA-MDF model proposed in this paper demonstrates significant advantages in both predictive performance and decision interpretability, providing a reference for formulating loan approval and risk prevention strategies.*

Keywords— CatBoost; Machine Learning; SHAP Framework; Interpretability.

I. INTRODUCTION

In recent years, as demand for social financing has continued to grow, loans have become an indispensable financial tool for both individuals and businesses. For financial institutions, the quality of loan approval decisions directly affects asset security and operational stability. Against this backdrop, credit scoring models serve as the foundation of automated approval systems by quantifying borrowers' credit risk and providing an objective basis for approval decisions [1]. A robust and accurate model not only improves approval efficiency and reduces default losses but also plays a key role in advancing the intelligent transformation of financial services.

However, traditional loan approval processes have long relied on manual experience, which comes with limitations such as low efficiency and inconsistent standards. While machine learning techniques have been widely adopted to build automated approval models, existing research still has gaps: on one hand, most models focus narrowly on default prediction and fail to adequately support comprehensive approval decisions involving multi-dimensional rules; on the other, many high-performance models are "black boxes" with opaque decision-making logic [2], making it difficult to meet the strict regulatory requirements for fairness and interpretability, and leaving applicants unable to understand the rationale behind approval outcomes. Therefore, exploring loan approval models that combine high predictive performance with strong interpretability has become an important focus for both academia and industry.

In response to these challenges, extensive research has been conducted in academia on loan approval modeling, forming three main directions.

In terms of constructing ensemble learning and complex models, Kokate and Chetty employed a combination of machine learning methods such as gradient boosting, random forests, and decision trees to build a credit scoring model for automated approval, using feature selection techniques to improve model efficiency. Their model demonstrated superior risk discrimination ability and stability on real-world banking data, validating the effectiveness of combining ensemble learning with feature optimization [3]. Further, Uddin et al. proposed a hybrid approach integrating deep learning with ExtraTrees and

adopted an ensemble voting mechanism to combine the three best-performing base models, achieving an accuracy of 87.26% in bank loan default prediction, highlighting the significant improvement of ensemble strategies on classification performance [4]. In line with this, Perera and Premaratne developed a stacking ensemble model based on a voting mechanism, further confirming the practicality and stability of ensemble learning in credit risk assessment [5]. Lakshmi and Rao constructed a loan default prediction model by integrating algorithms such as naive Bayes, decision trees, and multi-layer perceptrons, training on Kaggle historical data and ultimately achieving 90% accuracy, further demonstrating that ensemble learning methods can effectively support banks in making loan approval decisions [6].

In terms of benchmarking and optimization of traditional machine learning models, a substantial body of research has focused on providing benchmarks for model selection through systematic comparisons. Singh employed logistic regression, random forests, and support vector machines to model loan approval, achieving an accuracy of 78.785% on historical data [7]. Nureni and Adekola systematically compared eight machine learning algorithms and found that logistic regression performed particularly well in terms of accuracy (up to 83.24%) and sensitivity (up to 97.76%) [8]. Tumuluru et al. pointed out that among models for default prediction based on historical customer data, the random forest algorithm achieved the best accuracy [9]. Viswanatha conducted a comprehensive comparison of algorithms including random forests, naive Bayes, decision trees, and KNN, with naive Bayes achieving the highest accuracy at 83.73% [10]. A systematic comparison by Sarkar et al. also indicated that random forest and logistic regression models slightly outperformed other compared algorithms in terms of overall performance. These extensive comparative studies provide empirical support for model selection across different scenarios [11].

In terms of model efficiency, system deployment, and interpretability, relevant research has shown a clear trajectory of development—starting from basic system construction, moving toward feature optimization and model comparison, and ultimately deepening into interpretability and ethical considerations. Gomathy et al. built an automated loan prediction system based on decision trees, aiming to reduce manual review errors and improve overall approval efficiency [12]. Karthikeyan and Ravikumar introduced the Boruta feature selection method combined with random forest models to screen key features, effectively enhancing the model's discriminative power [13]. Bhattad et al. integrated multiple machine learning algorithms into a loan prediction system to improve the responsiveness and accuracy of the approval process [14]. Nalawade et al. further incorporated multi-algorithm comparison in system development, finding that logistic regression achieved the highest accuracy (88.70%), and designed a user interface while also noting the limitations of such models in practice due to their reliance on multi-attribute comprehensive judgment [15]. Subsequently, studies by Sharma et al. and Krishnaraj et al. both emphasized the potential of logistic regression in balancing predictive accuracy with model interpretability [16, 17]. More recent research has placed greater emphasis on the synergy between system performance and ethical standards: Badhan et al. built an approval system with high accuracy and anti-fraud capabilities through algorithm comparison and ensemble techniques [18]; Raheem further pointed out that while improving model performance, attention must also be paid to decision transparency and ethical considerations, thereby offering an important perspective for the sustainable development of loan decision systems [19].

Looking at the body of research discussed above, existing studies have laid a solid foundation for loan approval identification, yet there remain two issues that warrant further exploration. First, most approaches rely on a single type of data and fail to effectively integrate multi-dimensional applicant information, which limits the discriminative power and robustness of the models. Second, current research generally overlooks model interpretability—an aspect that is critically important for regulatory compliance and risk control in the context of loan approval identification.

To address these issues, this paper focuses on the task of loan approval identification and proposes an Interpretable Loan Approval Identification Model Using Multi-Dimensional Features (ILA-MDF). This study differs from existing research in the following two aspects. First, at the identification level, it builds a high-precision model for approval outcome prediction that outperforms existing benchmark models in classification performance while demonstrating strong generalization ability. Second, at the interpretability level, it introduces the SHAP framework to reveal the influence of key features on model decisions, significantly enhancing model interpretability and providing a reference for the formulation of loan approval risk management strategies.

II. MODEL AND THEORY

2.1 ILA-MDF Framework:

This study focuses on the issue of loan approval identification and divides the overall modeling process into three core stages: data source and processing, model building, and interpretability analysis, as illustrated in Figure 1.

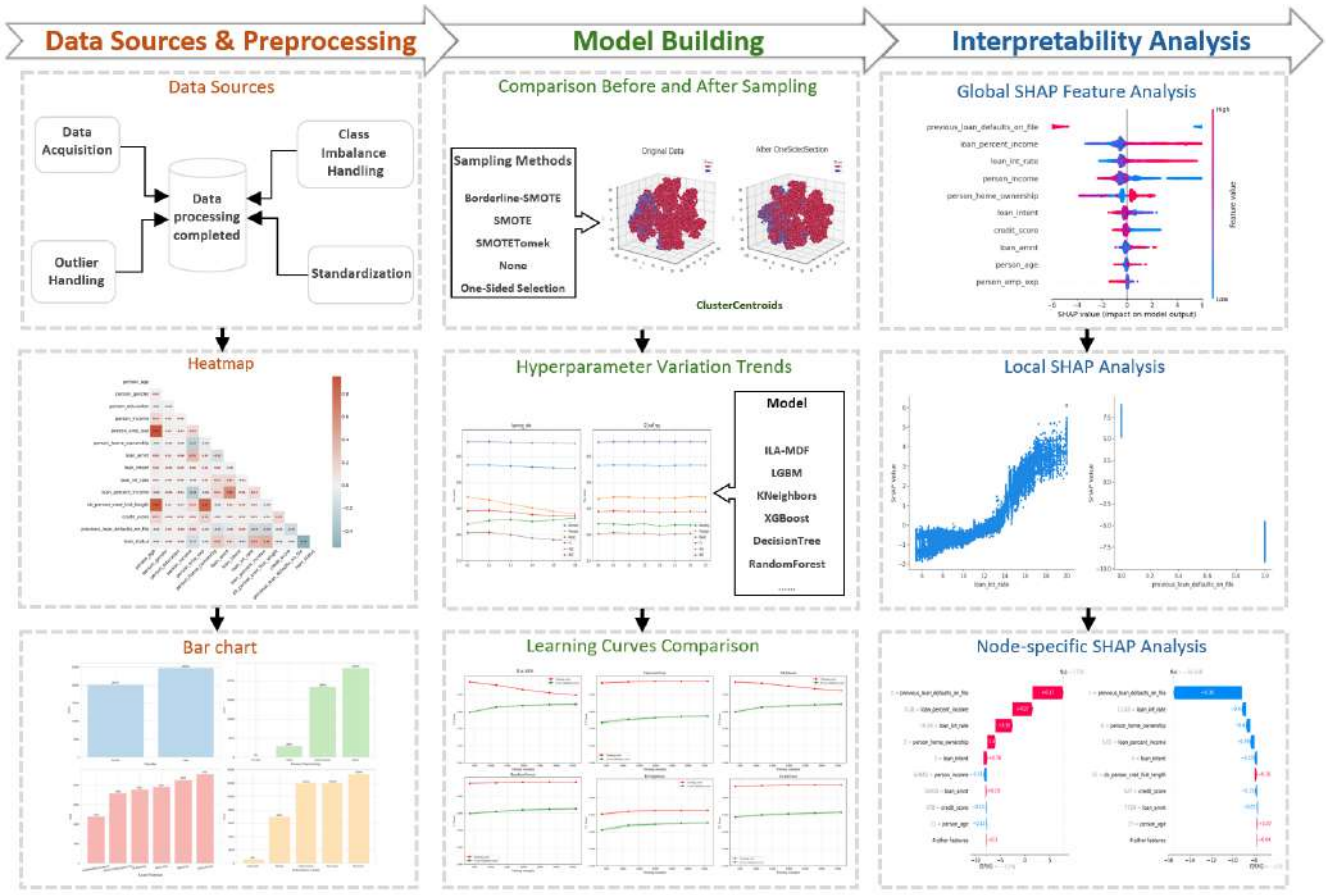


FIGURE 1: ILA-MDF Modeling Flowchart

Figure 1 shows the three-stage framework of the proposed ILA-MDF model. Stage 1 (Data Source and Processing) involves outlier detection, missing value handling, standardization, and One-Sided Selection undersampling. Stage 2 (Model Building) involves CatBoost model construction, grid search with cross-validation for hyperparameter tuning, and comprehensive performance evaluation. Stage 3 (Interpretability Analysis) involves SHAP-based global feature importance analysis, dependence plots for feature impact analysis, and waterfall plots for individual sample decision path analysis.*

1) **Data Source and Processing:** The data is sourced from the Kaggle platform, consisting of 45,000 records with 14 features. In terms of data preprocessing, outlier and missing value detection are first performed to ensure data integrity, followed by standardization. Subsequently, to address class imbalance, One-Sided Selection is introduced for undersampling to improve the identification performance of minority class instances.

2) **Model Building:** This paper constructs the ILA-MDF model based on CatBoost and evaluates its performance comprehensively across six dimensions: accuracy, precision, recall, F1 score, AUC, and Matthews correlation coefficient. On this basis, grid search combined with five-fold cross-validation is further employed to fine-tune the model's hyperparameters, enhancing its generalization ability and robustness.

3) **Interpretability Analysis:** This paper introduces the SHAP interpretability framework to analyze the model's decision mechanism from both global and local perspectives. Specifically, a summary plot is used to reveal the importance ranking of features in predicting outcomes; dependence plots focus on representative features, illustrating the relationship between their value variations and SHAP values; furthermore, waterfall plots are employed to analyze the prediction process of individual positive and negative samples, depicting the contribution and direction of influence of features on the final outcome.

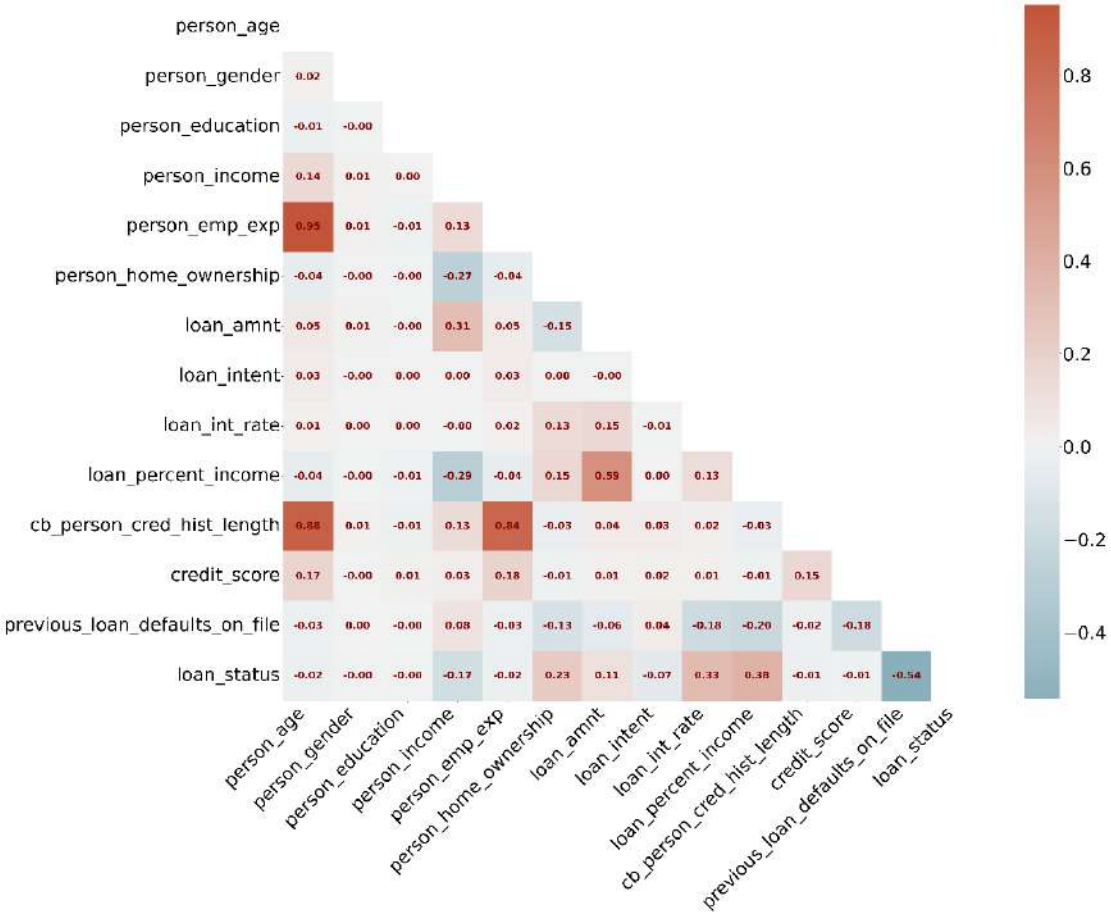


FIGURE 2: Correlation among Features

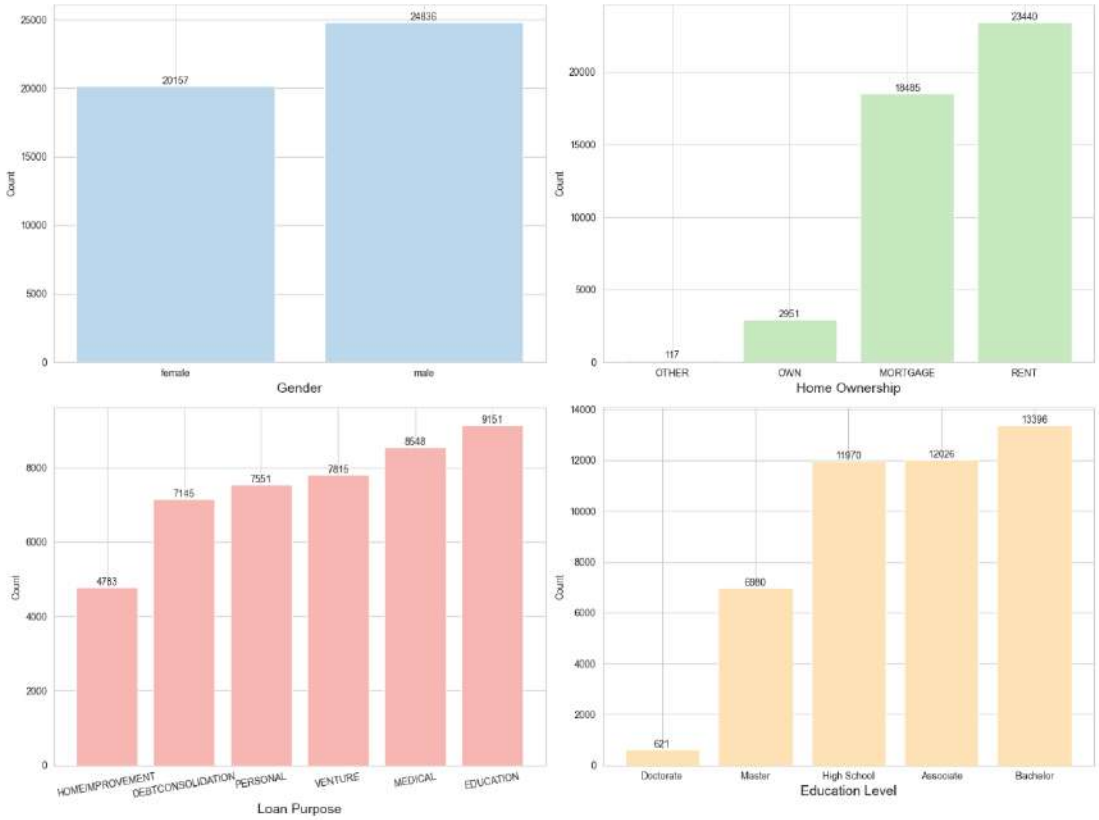


FIGURE 3: Proportion of Categorical Variables

2.2 CatBoost Algorithm:

This paper constructs the ILA-MDF model based on CatBoost. CatBoost is an open-source machine learning algorithm based on gradient boosting decision trees, proposed by Prokhorenkova et al. [20] in 2018. Compared with traditional GBDT, CatBoost introduces two key improvements: first, through "ordered target statistics" and "ordered boosting" mechanisms, it effectively alleviates target leakage in categorical feature encoding and reduces gradient estimation bias; second, it uses "symmetric trees" as base learners, which not only improves inference efficiency and generalization performance but also enhances parallelization in storage and computation. Based on these mechanisms, this paper further divides the derivation and optimization of the objective function into the following three steps:

2.2.1 Objective Function Definition:

The objective function consists of a loss function and a regularization term, as shown in Equation (1):

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{1}$$

where l is the loss function, $\hat{y}_i^{(t-1)}$ is the predicted value from the previous $t-1$ iterations, $f_t(x_i)$ is the tree added at the t-th iteration, and $\Omega(f_t)$ is the regularization term. Its specific form is:

$$\Omega(f) = \gamma T + \lambda \sum_{j=1}^{T} w_j^2 \tag{2}$$

where T is the number of leaf nodes, γ controls the penalty on the number of leaves, λ is the L2 regularization coefficient, and w_j is the weight of leaf node j.

2.2.2 Ordered Boosting Strategy:

To avoid target leakage, CatBoost introduces an ordered boosting strategy in gradient computation. Specifically, for a sample x_i, its gradient depends only on historical information from the first $i-1$ samples:

$$g_i = \frac{\partial l(y_i, \hat{y}^{(t-1)})}{\partial \hat{y}^{(t-1)}} \tag{3}$$

$$h_i = \frac{\partial^2 l(y_i, \hat{y}^{(t-1)})}{\partial (\hat{y}^{(t-1)})^2} \tag{4}$$

2.2.3 Symmetric Tree Design and Leaf Weight Optimization:

CatBoost adopts a symmetric tree structure, where nodes at the same level share splitting rules, thereby improving computational efficiency while ensuring model stability. For leaf node j, the optimal leaf weight w_j^* can be obtained by minimizing the objective function:

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{5}$$

where I_j is the set of samples in leaf node j. Let $G_j = \sum_{i \in I_j} g_i$ and $H_j = \sum_{i \in I_j} h_i$. Substituting these into the objective function yields:

$$Obj^{(t)} = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{6}$$

In summary, CatBoost automatically adjusts category weights to improve the identification of minority class samples in loan approval identification. Its gradient computation mechanism based on historical samples effectively suppresses target leakage and reduces prediction bias. The symmetric tree structure enhances generalization performance and inference efficiency while controlling overfitting.

2.3 SHAP Framework:

SHAP (SHapley Additive exPlanations) is a unified framework proposed by Lundberg et al. [21] in 2017 for explaining predictions made by black-box machine learning models. The framework is based on the concept of Shapley values from cooperative game theory, and it converts the output of black-box models into interpretable numerical results by fairly distributing the marginal contribution of each feature to the prediction, thereby meeting interpretability requirements while preserving predictive performance. The formula for SHAP values is as follows:

$$f(x_i) = f_0 + f(x_{i,1}) + f(x_{i,2}) + \cdots + f(x_{i,k}) \quad (7)$$

where $x_{i,k}$ represents the k-th feature of the i-th sample; $f(x_{i,k})$ represents the Shapley value of $x_{i,k}$; f_0 represents the baseline prediction of the overall model; and $f(x_i)$ represents the predicted value for the i-th sample. When $f(x_{i,k}) > 0$, the feature has a positive contribution to the prediction; when $f(x_{i,k}) < 0$, the feature has a negative contribution. Therefore, introducing the SHAP framework in this study not only enables the identification of key features influencing classification outcomes in loan approval based on the magnitude of feature SHAP values, helping to pinpoint potential risk factors, but also provides deeper insight into the model's decision-making mechanisms and working principles, offering a reliable basis for model optimization.

III. RESULTS AND DISCUSSION

3.1 Data Source:

The dataset used in this study is sourced from the Kaggle platform, consisting of 45,000 records with 14 features. The specific details are presented in Table 1. In terms of data preprocessing, after data inspection, outliers were identified and removed, and categorical variables were one-hot encoded, resulting in a complete dataset that laid the foundation for subsequent model training and analysis.

TABLE 1
DESCRIPTION OF DATASET FEATURES

Feature	English explanation
person_age	Age
person_gender	Gender
person_education	Highest Education Level
person_income	Annual Income
person_emp_exp	Years of Work Experience
person_home_ownership	Home Ownership Status
loan_amnt	Loan Amount Applied
loan_intent	Loan Purpose
loan_int_rate	Loan Interest Rate
loan_percent_income	Loan Amount to Annual Income Percentage
cb_person_cred_hist_length	Credit History Length
credit_score	Credit Score
previous_loan_defaults_on_file	Previous Loan Default Indicator
loan_status	Loan Approval Status

3.2 Feature Correlation Analysis:

To gain deeper insight into the relationship between various features and loan approval status, this study conducted a systematic analysis of linear correlations among variables using a heatmap. The results reveal the differentiated roles that different features play in the approval decision.

Specifically, the ratio of loan amount to annual income (loan_percent_income) shows a moderate positive correlation with approval status (r = 0.38), indicating that this metric serves as an important predictor of approval outcomes—that is, the higher the debt-to-income ratio, the lower the likelihood of loan approval. A stronger negative correlation was observed for previous loan defaults on file (r = -0.54), further confirming the significant impact of this feature on approval decisions: applicants with

a history of default face a considerably higher risk of rejection. In addition, home ownership status (person_home_ownership) exhibits a certain positive correlation with approval status (r = 0.23), suggesting that applicants who own a home are generally perceived as having a more stable credit background and are therefore more likely to be approved.

It is worth noting that credit score shows no significant linear correlation with approval status (r = -0.01), indicating that, within the current modeling framework, this variable offers limited explanatory power for approval outcomes. Its potential influence may need to be further explored through nonlinear modeling approaches.

3.3 Analysis of Categorical Variables:

The dataset includes four main categorical variables: gender, home ownership, loan purpose, and default history. Gender is roughly balanced, reducing the risk of bias. Home ownership is concentrated among renters and mortgage holders, with owners and other categories making up only a small share—a pattern that reflects the link between housing status and financial stability. Loan purpose covers a range of categories such as education, personal consumption, and medical expenses, capturing diverse real-world borrowing scenarios. Default history, however, is highly imbalanced: only a small fraction of applicants have prior defaults, which calls for sampling techniques in modeling to improve identification of this minority but critical group.

3.4 Experimental Environment and Evaluation Metrics:

The experiments were conducted in the following software and hardware environment: Intel(R) Core(TM) i5-10200H processor (base frequency 2.40 GHz), 16GB RAM, and Windows 10 operating system. The development environment was PyCharm 2023.2 with Python 3.11.

TABLE 2
CONFUSION MATRIX

	Actual Positive	**Actual Negative**
Predicted Positive	TP (True Positive)	FP (False Positive)
Predicted Positive	TP (True Positive)	FP (False Positive)

Six metrics were selected to evaluate model performance: accuracy, precision, recall, F1 score, AUC, and Matthews correlation coefficient (MCC). These metrics were calculated based on the confusion matrix by comparing predicted results with actual labels.

Accuracy measures the proportion of samples correctly predicted by the model:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Precision focuses on how many of the samples predicted as positive are actually positive:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{9}$$

Recall measures how many of the actual positive samples are correctly identified:

$$\text{Recall} = \frac{TP}{TP+FN}) \tag{10}$$

The F1 score is the harmonic mean of precision and recall:

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}} = \frac{2TP}{2TP+FP+FN} \tag{11}$$

AUC evaluates the overall performance of the model by plotting the true positive rate (TPR) against the false positive rate (FPR); the larger the area, the stronger the model's ability to distinguish between positive and negative classes.

$$TPR = \frac{TP}{TP+FN}, FPR = \frac{FP}{FP+TN} \tag{12}$$

MCC takes into account all elements of the confusion matrix (TP, TN, FP, FN) and provides a comprehensive measure of classification quality.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{13}$$

3.5 Imbalance Handling:

The dataset in this study suffers from class imbalance. To mitigate its impact on model training, the One-Sided Selection method was applied for sampling. Figure 4 shows a comparison of the data distribution before and after sampling. The left panel shows a clear class imbalance in the original data, with samples of class 0 significantly outnumbering those of class 1. The right panel shows that after sampling, the two classes become more balanced, which helps improve the model's ability to identify the minority class.

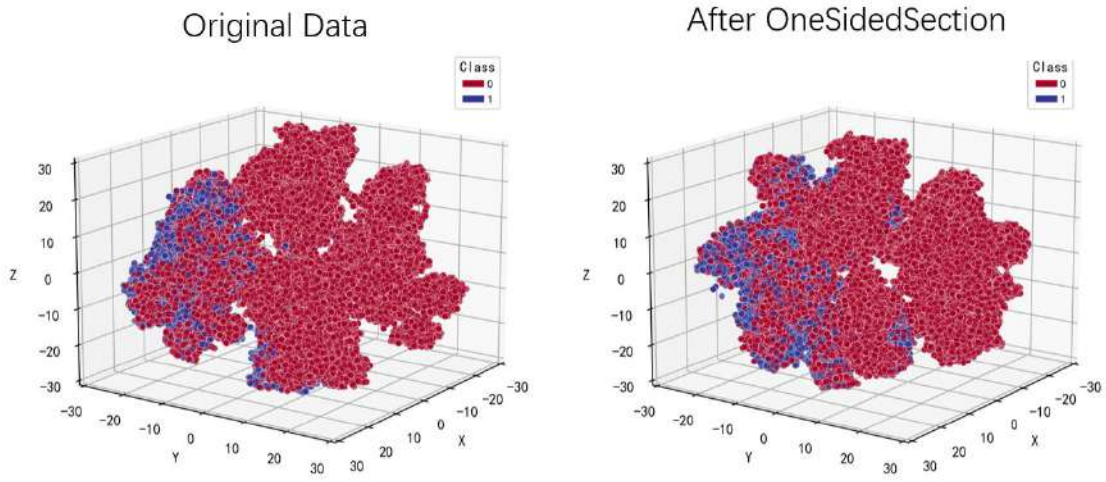


FIGURE 4: Comparison Before and After Sampling

As shown in Figure 4, the One-Sided Selection sampling method effectively alleviates the class imbalance issue. To further validate its impact on model performance, this study compares different sampling methods experimentally.

The results are shown in Table 3. The One-Sided Selection method performs best across three key metrics: recall, F1 score, and MCC. In subsequent experiments, One-Sided Selection is adopted as the default sampling strategy for further modeling and evaluation.

TABLE 3
COMPARISON OF IMBALANCE HANDLING RESULTS

Sampling	Accuracy	Precision	Recall	F1	AUC	MCC
Borderline-SMOTE	93.22%	87.46%	81.15%	84.18%	97.75%	79.97%
SMOTE	93.26%	87.67%	81.10%	84.25%	97.77%	80.08%
SMOTETomek	93.28%	87.82%	81.02%	84.28%	97.78%	80.12%
None	93.36%	89.80%	79.12%	84.12%	97.81%	80.20%
One-Sided Selection	93.30%	87.23%	81.84%	84.45%	97.77%	80.25%

3.6 Comparison with Mainstream Machine Learning Models:

To comprehensively evaluate the performance of the ILA-MDF model, this study selects a variety of mainstream machine learning models as baseline methods for comparative analysis. All models are trained and tested under the same feature engineering and data preprocessing pipeline, with evaluation metrics including accuracy, precision, recall, F1 score, AUC, and MCC.

As shown in Table 4, ILA-MDF ranks first in accuracy, precision, F1 score, and MCC, demonstrating strong classification performance and stability. Although XGBoost achieves higher recall and AUC, ILA-MDF maintains a clear advantage in overall performance. With an accuracy of 93.30%, ILA-MDF exhibits excellent classification capability. In contrast, traditional models like KNeighbors, while acceptable in accuracy, often underperform in recall and F1 when handling nonlinear, high-dimensional, and imbalanced data, limiting their applicability in complex real-world scenarios. Overall, ILA-MDF performs well across multiple metrics, maintaining high predictive accuracy and generalization ability, particularly in high-dimensional imbalanced settings.

TABLE 4
COMPARISON WITH MAINSTREAM MACHINE LEARNING MODELS

Model	Accuracy	Precision	Recall	F1	AUC	MCC
KNeighbors	89.06%	75.49%	75.18%	75.33%	92.73%	68.31%
LogisticRegression	88.89%	74.22%	76.67%	75.42%	94.95%	68.27%
DecisionTree	89.70%	75.28%	79.88%	77.51%	86.19%	70.89%
GradientBoosting	91.99%	84.44%	78.39%	81.30%	97.02%	76.29%
ExtraTrees	91.94%	83.59%	79.33%	81.40%	96.92%	76.31%
RandomForest	92.58%	86.27%	79.22%	82.59%	97.30%	78.00%
HistGradientBoosting	93.06%	86.61%	81.36%	83.90%	97.68%	79.55%
LGBM	93.09%	86.62%	81.49%	83.97%	97.70%	79.64%
XGBoost	93.21%	86.08%	82.82%	84.42%	97.79%	80.10%
ILA-MDF	93.30%	87.23%	81.84%	84.45%	97.77%	80.25%

3.7 Comparison with Existing Studies:

To highlight the performance advantages of the ILA-MDF model in loan approval classification tasks, this study compares it with representative models from existing literature. The comparison covers traditional machine learning methods, deep learning models, and mainstream ensemble algorithms to ensure the validity of the results.

As shown in Table 5, the ILA-MDF model achieves the best performance across multiple key metrics, including accuracy, precision, F1 score, AUC, and MCC, significantly outperforming all benchmark models. It demonstrates not only stable classification capability but also strong generalization in handling imbalanced data and complex pattern recognition. Therefore, the model not only outperforms existing methods in a statistically significant manner but also shows strong potential and robustness for real-world applications.

TABLE 5
COMPARISON WITH EXISTING STUDIES

Author	Model	Accuracy	Precision	Recall	F1	AUC	MCC
Uddin et al.[4]	Ensemble	82.76%	/	/	/	/	/
Perera et al. [5]	Stacking	78.00%	78.00%	92.00%	84.00%	75.00%	/
Nureni et al. [8]	LR	78.16%	79.00%	75.96%	75.96%	/	/
Natasha et al. [22]	DNN	/	/	/	/	63.80%	/
Shinde et al. [23]	LR	72.14%	/	/	82.79%	/	/
Ndayisenga [24]	GB	81.11%	82.14%	80.98%	81.47%	/	/
Yang [25]	LR	76.00%	/	/	/	/	/
Ours	ILA-MDF	93.30%	87.23%	81.84%	84.45%	97.77%	80.25%

3.8 Hyperparameter Perturbation Analysis:

To enhance the robustness and generalization ability of the model, a grid search strategy was employed for hyperparameter optimization. The tuning focused on four key parameters, with the F1 score used as the performance evaluation metric. The search range and final optimal values for each parameter are shown in Table 6.

TABLE 6
HYPERPARAMETER TUNING RESULTS

Parameter	Search Range	Optimal Value
Iterations	[300, 400, 500]	500
Learning_rate	[0.01, 0.05, 0.1]	0.1
Depth	[4, 6, 8]	6
L2_leaf_reg	[0.5, 1, 1.5]	0.5

After training the model with the optimal hyperparameter combination, the performance on the test set was compared before and after tuning. The results are shown in Table 7. All metrics improved after tuning. Specifically, recall increased by 0.19%, F1 score by 0.08%, and MCC by 0.09%, indicating that hyperparameter optimization effectively enhanced model classification performance and stability.

TABLE 7
MODEL PERFORMANCE BEFORE AND AFTER TUNING

	Accuracy	Precision	Recall	F1	AUC	MCC
Before Tuning	93.30%	87.23%	81.84%	84.45%	97.77%	80.25%
After Tuning	93.33%	87.19%	82.03%	84.53%	97.77%	80.34%

To further analyze the impact of hyperparameters on model performance, this study examined the effect of varying the learning rate and L2 regularization coefficient on key metrics (see Figure 5). The left panel shows that when the learning rate is set to 0.1, all metrics reach their highest values, and as the learning rate continues to increase, performance declines across the board, indicating that an excessively high learning rate suppresses the model's learning ability and leads to underfitting. Therefore, the learning rate was ultimately set to 0.1.

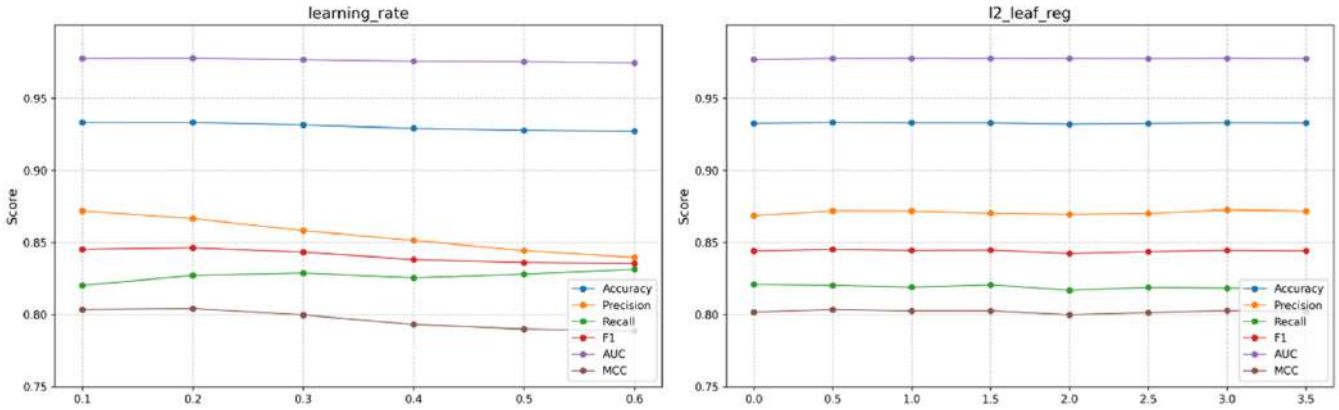


FIGURE 5: Hyperparameter Variation Trend Chart

The right panel illustrates the effect of the L2 regularization coefficient on model performance. The results show that as the coefficient increases, all metrics except recall show an upward trend at 0.5. However, beyond 0.5, precision, recall, F1, and MCC exhibit a fluctuating pattern of initial decline followed by a subsequent rise. Considering overall performance, the optimal effect is achieved at a coefficient of 0.5, which was ultimately chosen.

3.9 SHAP Interpretability Analysis:

To enhance the interpretability of the model, this study introduces the SHAP framework, employing summary plots, dependence plots, and waterfall plots to visually analyze the model's decision-making process. The analysis explores the key

factors influencing loan approval identification from three perspectives: global feature contribution, feature impact mechanisms, and individual sample decision paths.

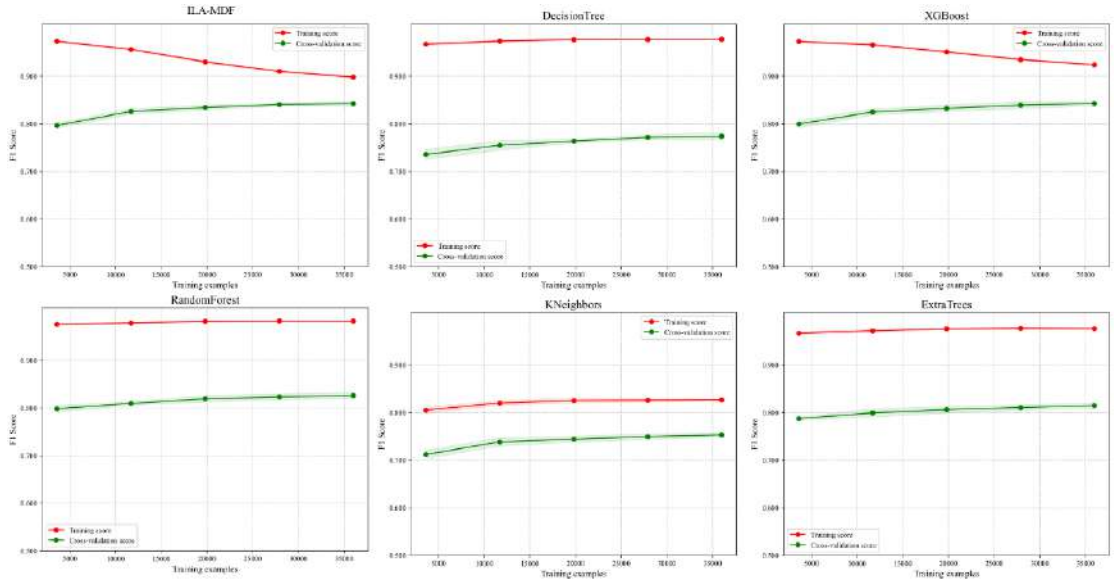


FIGURE 6: Comparison of Model Learning Curves

3.9.1 Global Feature Importance Analysis:

To examine the importance and direction of influence of each feature in the model's decisions, a global interpretability analysis was conducted using the SHAP method. Based on the SHAP summary plot shown in Figure 7, the following conclusions can be drawn:

The indicator of previous loan defaults is the most influential feature on model predictions. When its value is high, it corresponds to a consistently negative SHAP value, indicating that a borrower's history of default systematically reduces the likelihood of loan approval. This underscores the strong inhibitory effect of this feature on approval decisions.

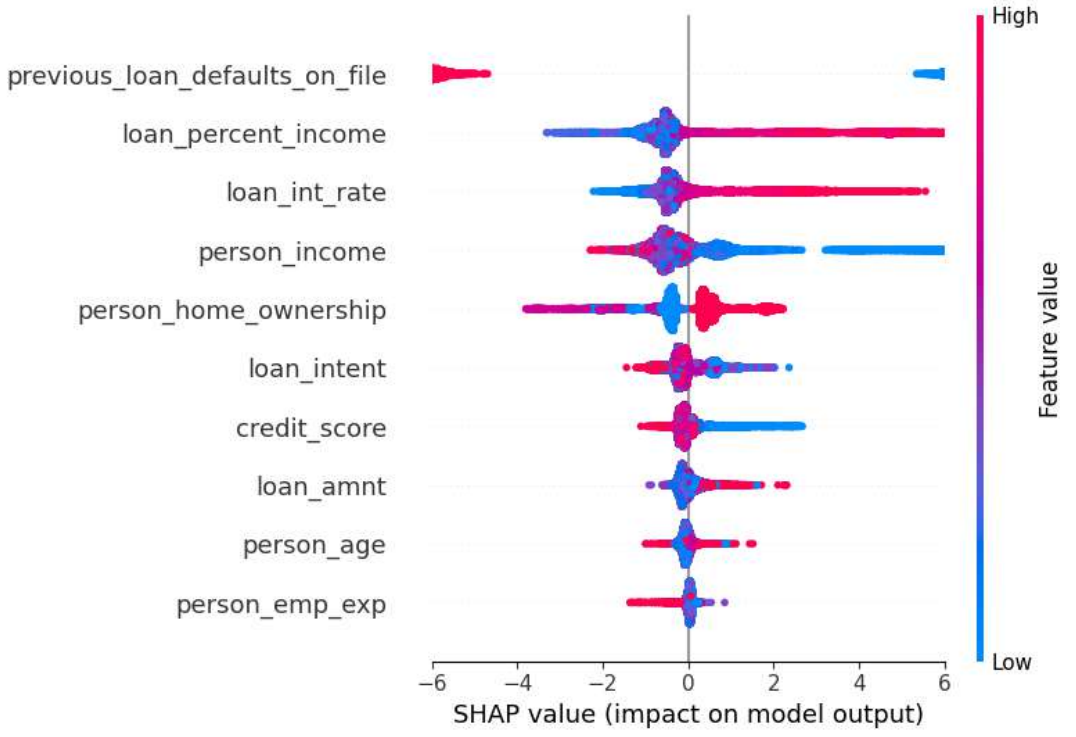


FIGURE 7: SHAP Summary Plot

When the ratio of loan amount to annual income is low, the SHAP value is clearly negative. This suggests that when the loan amount is relatively small compared to the borrower's annual income, the repayment burden is lighter, and the model tends to classify the applicant as "low risk," thereby significantly increasing the probability of approval.

Similarly, when the interest rate is low, the SHAP value is also noticeably negative. Since lower interest rates are typically offered to high-quality borrowers, a low rate itself serves as a signal of lower risk, positively influencing the final approval outcome.

In contrast, annual income generates a positive SHAP value when its value is low. This means that when a borrower's income level is low, this feature clearly pushes the model's prediction toward the "high-risk" direction, significantly increasing the likelihood of rejection. This reflects the model's sensitivity to insufficient repayment capacity.

3.9.2 Analysis of Feature Impact on Prediction Results:

To gain deeper insight into how key features influence model predictions, this study uses SHAP values to generate dependence plots (shown in Figure 8), revealing the relationship between feature values and their corresponding contributions.

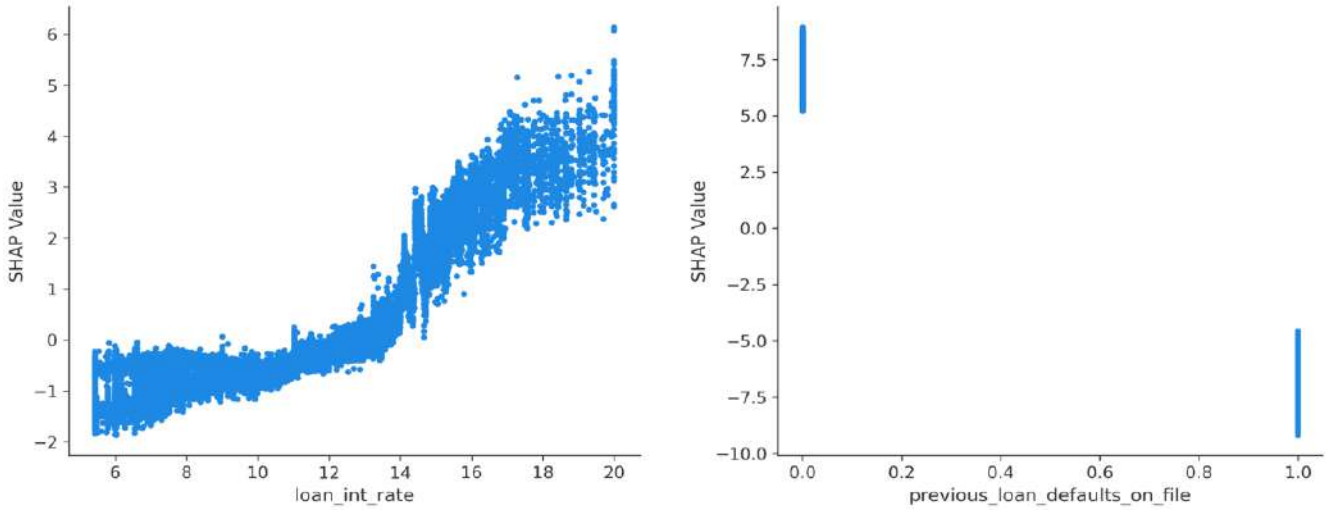


FIGURE 8: SHAP Dependence Plot

The results show that the loan interest rate (loan_int_rate) exhibits a significant positive correlation with model risk prediction. When the interest rate exceeds 14%, the SHAP value begins to rise rapidly, forming a stable positive driving effect in the range of 17% to 20%. This indicates that a higher interest rate significantly increases the probability of the model classifying an applicant as high-risk. This aligns with real-world practice: the higher the loan rate, the more inclined financial institutions are to approve the loan application.

For the indicator of previous loan defaults (previous_loan_defaults_on_file), when the value is 0—indicating no history of default—the corresponding SHAP value is positive, reflecting that having no default record has a positive effect on loan approval. Conversely, when a default exists, the SHAP value is negative, indicating that a record of default negatively influences the likelihood of loan approval.

3.9.3 Analysis of Sample Decision Path:

To better understand the model's decision-making mechanism at the sample level, this study further examines how individual features collectively contribute to the final prediction from a single-sample perspective. The SHAP waterfall plot shown in Figure 9 visually illustrates the cumulative contribution of each feature to the prediction for a given sample.

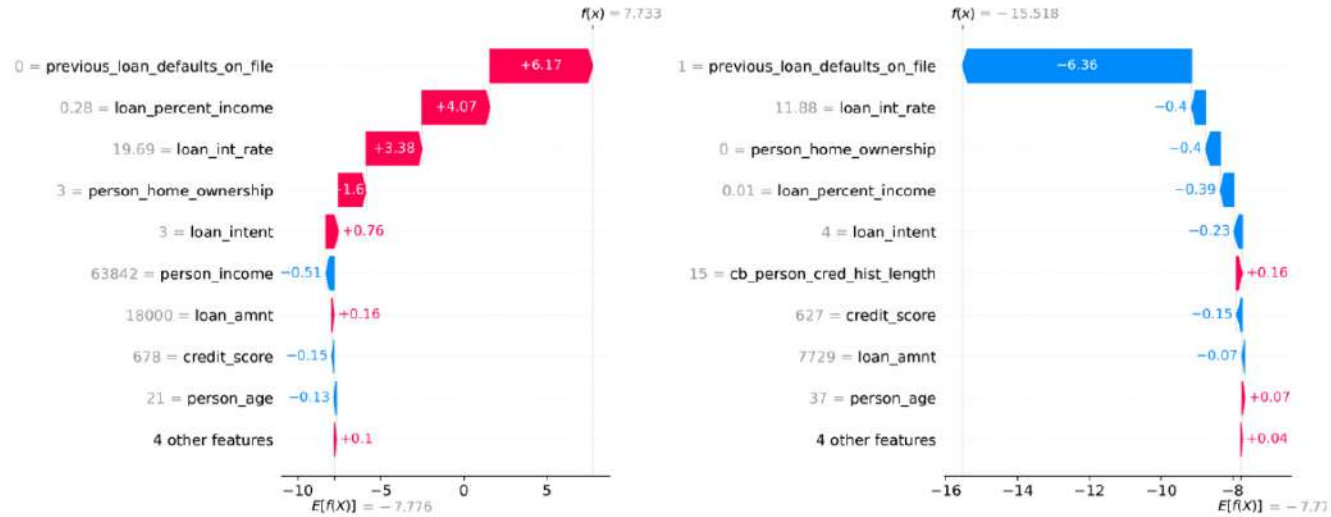


FIGURE 9: SHAP Waterfall Plot

In the approved sample (left panel), the indicator of previous loan defaults is the most significant positive contributor, indicating that the applicant's clean credit history plays a key role in the approval decision. The loan-to-income ratio and loan interest rate also show negative contributions, suggesting that a lower debt burden and a more favorable interest rate together point to the borrower's strong repayment capacity and low credit risk, thereby significantly increasing the likelihood of approval.

In contrast, for the rejected sample (right panel), the loan-to-income ratio shows a strong negative contribution, indicating that a high debt burden is one of the main reasons the model denies the application. Meanwhile, annual income also exhibits a clear negative influence, further showing that low income combined with high debt pressure reinforces the applicant's credit risk, ultimately leading the model to make a rejection decision.

IV. CONCLUSION

This paper focuses on the problem of loan approval identification and proposes a model named ILA-MDF (Interpretable Loan Approval Identification Model based on Multidimensional Features).

First, outlier detection is performed on the acquired dataset, followed by standardization and One-Sided Selection undersampling, effectively improving data quality and class balance. One-Sided Selection was chosen for its ability to remove noisy majority class samples while preserving informative decision boundary samples.

Second, through a comprehensive comparison of multiple mainstream machine learning algorithms, CatBoost is selected as the base architecture, and grid search combined with five-fold cross-validation is used for hyperparameter tuning, significantly enhancing the model's generalization ability and robustness.

Third, multi-level analysis based on SHAP shows that features such as previous loan defaults, the ratio of loan amount to annual income, and loan interest rate have a significant impact on identification results. In particular, when the loan interest rate exceeds 14% or the indicator for previous loan defaults is 0, the probability of loan approval increases significantly. These findings provide a theoretical basis for risk control strategies in loan approval.

LIMITATIONS AND FUTURE WORK

This study has several limitations that should be acknowledged. First, the dataset used is cross-sectional, meaning the model does not account for temporal dynamics in borrower behavior or changing economic conditions. Second, while the model demonstrates strong predictive performance, potential fairness concerns—such as disparate impact across demographic groups—were not evaluated. Third, the computational cost and inference time of the CatBoost model compared to simpler baselines were not analyzed, which may be relevant for real-time deployment scenarios.

Future work will focus on addressing these limitations. Specifically, we aim to incorporate users' historical credit behavior data to construct time-aware models, enabling loan approval decisions to evolve from static assessments toward continuous monitoring and early warning based on dynamic risk profiling. Additionally, we plan to integrate fairness-aware machine

learning techniques to ensure equitable outcomes across diverse applicant groups and conduct computational efficiency analyses to support real-world deployment.

ACKNOWLEDGEMENTS

The author has no acknowledgements to declare.

CONFLICT OF INTEREST

The author declares no conflict of interest

CONFLICT OF INTEREST

The authors have no acknowledgements to declare.

REFERENCES

[1] Sarathamani, S., & Kumar, S. (2024). Application of machine learning in credit scoring for loan approval systems. In *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1–7). IEEE.

[2] Zou, Y., Xia, M., & Lan, X. (2025). Interpretable credit scoring based on an additive extreme gradient boosting. *Chaos, Solitons & Fractals, 194*, Article 116216.

[3] Kokate, S., & Chetty, M. S. R. (2021). Credit risk assessment of loan defaulters in commercial banks using voting classifier ensemble learner machine learning model. *International Journal of Safety and Security Engineering, 11*(5), 565–572.

[4] Uddin, N., Ahamed, M. K. U., Uddin, M. A., Memon, M. H., & Uddin, M. A. (2023). An ensemble machine learning based bank loan approval predictions system with a smart application. *International Journal of Cognitive Computing in Engineering, 4*, 327–339.

[5] Perera, C. L., & Premaratne, S. C. (2024). An ensemble machine learning approach for forecasting credit risk of loan applications. *WSEAS Transactions on Systems, 23*, 31–46.

[6] Lakshmi, M. V. N., & Rao, P. S. (2024). A prediction of modernized loan approval system based on machine learning approach. *International Journal for Modern Trends in Science and Technology, 10*(6), 17–21.

[7] Singh, V. (2021). Prediction of modernized loan approval system based on machine learning approach. In *2021 International Conference on Intelligent Technologies (CONIT)* (pp. 1–6). IEEE.

[8] Nureni, A. A., & Adekola, O. E. (2022). Loan approval prediction based on machine learning approach. *FUDMA Journal of Sciences, 6*(3), 41–50.

[9] Tumuluru, P., Burra, L. R., Loukya, M., Sai, N. S., & Bhavani, K. (2022). Comparative analysis of customer loan approval prediction using machine learning algorithms. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (pp. 349–353). IEEE.

[10] Viswanatha, V., Ramachandra, A. C., Vishwas, K. N., Kavyashree, B. S., & Nayana, G. (2023). Prediction of loan approval in banks using machine learning approach. *International Journal of Engineering and Management Research, 13*(4), 7–19.

[11] Sarkar, T., Rakhra, M., Sharma, V., Singh, S., & Sharma, P. (2024). An empirical comparison of machine learning techniques for bank loan approval prediction. In *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)* (pp. 137–143). IEEE.

[12] Gomathy, C. K., Charulatha, M., Aakash, M., & Mani, P. (2021). The loan prediction using machine learning. *International Research Journal of Engineering and Technology, 8*(10), 2395–0056.

[13] Karthikeyan, S. M., & Ravikumar, P. (2021). A comparative analysis of feature selection for loan prediction model. *International Journal of Computer Applications, 975*, 8887.

[14] Bhattad, S., Bawane, S., Agrawal, S., & Tiwari, A. (2021). Loan prediction using machine learning algorithms. *International Journal of Computer Science Trends and Technology, 9*(3), 143–146.

[15] Nalawade, S., Andhe, S., Parab, S., & Gurav, S. (2022). Loan approval prediction. *International Research Journal of Engineering and Technology, 9*(4), 669–673.

[16] Sharma, H., Tyagi, I., Agarwal, G., & Sharma, A. (2023). An exhaustive investigation on loan prediction in banks using LRD. *[Journal name missing]*.

[17] Krishnaraj, P., Rita, S., & Jaiswal, J. (2023). Comparing machine learning techniques for loan approval prediction. In *Proceedings of the 1st International Conference on Artificial Intelligence, Communication, IoT, Data Engineering and Security (IACIDS)* (pp. 23–25).

[18] Badhan, A., Rana, A., Malhi, S. S., & Singh, J. (2024). A comparative analysis for loan approval prediction using machine learning. In *2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT)* (Vol. 1, pp. 1–5). IEEE.

[19] Raheem, M. (2024). Loan default prediction using machine learning: A review on the techniques. *Journal of Applied Technology and Innovation, 8*(2), 1.

[20] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems, 31*, 6638–6648.

[21] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 4765–4774.

[22] Natasha, A., Prastyo, D. D., & Suhartono. (2019). Credit scoring to classify consumer loan using machine learning. In *AIP Conference Proceedings* (Vol. 2194, No. 1, Article 020070). AIP Publishing.

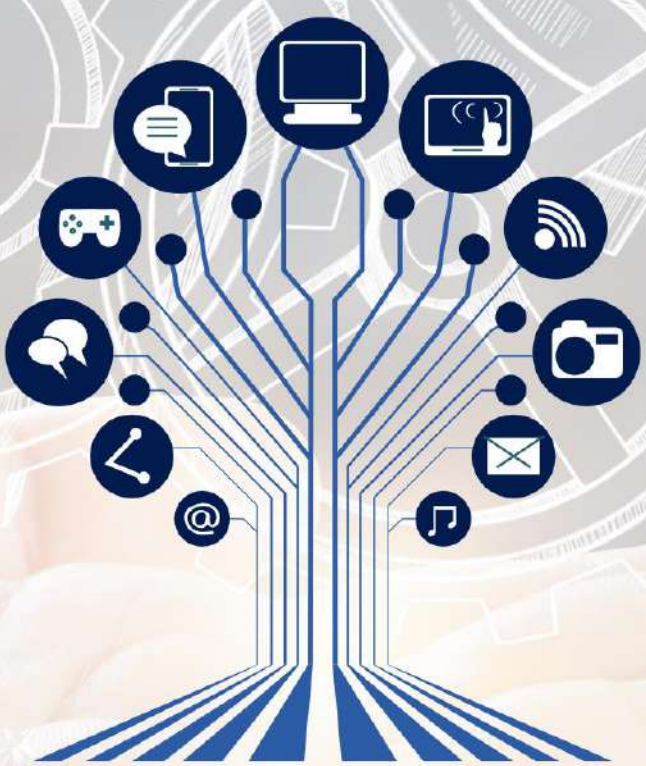
[23] Shinde, A., Patil, Y., Kotian, I., & Kharat, S. (2022). Loan prediction system using machine learning. In *ITM Web of Conferences* (Vol. 44, Article 03019). EDP Sciences.

[24] Ndayisenga, T. (2021). *Bank loan approval prediction using machine learning techniques* [Master's thesis, [University name]].

[25] Yang, C. (2024). Research on loan approval and credit risk based on the comparison of machine learning models. In *SHS Web of Conferences* (Vol. 181, Article 02003). EDP Sciences.

IJOER
ENGINEERING JOURNAL

International Journal of Engineering Research and Science



Published by

AD Publications

Contact us +91-7665235235 www.ijoer.com info@ijoer.com