ICBDT
Big Data Technologies

Conference Program

2026 8th International Conference on Big Data Technologies (ICBDT 2026)

May 22-24 | Nanchang, China

Sponsored by

南昌航空大学
NANCHANG HANGKONG UNIVERSITY

Organized by

南昌航空大学
信息工程学院

Supported by

南昌大学 NANCHANG UNIVERSITY
江西财经大学 JIANGXI UNIVERSITY OF FINANCE AND ECONOMICS
江西省城市安全信息化学会 Jiangxi Urban Security Information Society

TABLE OF CONTENTS

WELCOME MESSAGE

Dear Colleagues and Distinguished Guests,

On behalf of the organizing committees, we are delighted to extend our warmest welcome to the 2026 8th International Conference on Big Data Technologies (ICBDT 2026), taking place from May 22-24, 2026, in Nanchang, China.

As a premier international forum sponsored by Nanchang Hangkong University, China, ICBDT 2026 brings together leading researchers, practitioners, and industry experts to share cutting-edge advancements in big data architecture and infrastructure, data analytics and machine learning, security and privacy, industry applications, visualization and interaction, and data management and governance. The conference features inspiring keynote speeches, peer-reviewed paper presentations, and vibrant discussions on frontier technologies and real-world applications.

We are honored to host this event in Nanchang-a city renowned as the “Hero City” for its profound revolutionary heritage, vibrant technological innovation, and warm hospitality. We sincerely hope your participation will spark meaningful collaborations, inspire innovative ideas, and create lasting memories.

Thank you for joining us. We look forward to an engaging and successful conference!

ICBDT 2026

Conference Committee

ORGANIZING COMMITTEE

Advisory Committee -

Yuming Fang, Vice President of Jiangxi University of Finance and Economics, China

Wei Huang, Yichun University, China

Yibin Huang, Xinyu University, China

Bangshu Xiong, Nanchang Hangkong University, China

Hong Rao, Nanchang University, China

Chengying Mao, Jiangxi University of Finance and Economics, China

Zichen Xu, Nanchang University, China

General Chairs

Dai Jiyang, Nanchang Hangkong University, China

William Wei Song, Dalarna University, Sweden

Program Chairs

Roger Nyberg, Dalarna University, Sweden

Qiaofeng Ou, Nanchang Hangkong University, China

Lan Yang, California State Polytechnic University, USA

Publicity Chairs

Yang Cihui, Nanchang Hangkong University, China

Fairouz Kamareddine, Heriot-Watt University, UK

Fabrizio Marozzo, University of Calabria, Italy

Local Organizing Chair

Sheng Zhang, Nanchang Hangkong University, China

Technical Committe

Ali Yavari, Swinburne University of Technology, Australia
Bingxian Ma, University of Jinan, China
Chen Jinpeng, Beijing University of Posts and Telecommunications, China
Cong Liu, Shandong University of Technology, China
Cun JI, Shandong Normal University, China
Daqiang Zhang, Tongji University, China
Gabriela Mogos, Xi'an Jiaotong-Liverpool University, China
Hui Zhang, Southwest University of Science and Technology, China

CONFERENCE VENUE

Nanchang Qianhu Hotel

(南昌前湖酒店)

Address: 616, Xue Fu Avenue,Hong Gu Tan, Nanchang,Jiangxi, China

TRAFFIC INFORMATION

➤ Departing from Nanchang Changbei International Airport

Taxi/Online Car-hailing: The driving distance is about 29–30 kilometers, taking approximately 30–40 minutes, and the cost is about 95–100 yuan.

Airport Shuttle Bus: Take Airport Bus Line 3 to Nanchang University Qianhu Campus Station or Qianhu Avenue North Intersection Station (about 35–40 minutes, 15 yuan), then walk about 650 meters to 1 kilometer to the hotel. The last bus departs around 21:30. Alternatively, take Airport Bus Line 2 to Fenghe Overpass Station, transfer to High-speed Rail Bus Line 1 or Bus 225, get off at the station near Jiangxi Science and Technology Normal University or Qingyuan Road North Intersection, and walk about 315 meters. The whole journey takes about one hour and costs about 16 yuan. Note that the last Bus 225 departs at 19:30.

➤ Departing from Nanchang West Railway Station

Taxi/Online Car-hailing: The driving distance is about 5–6 kilometers, taking approximately 10–15 minutes, and the cost is about 15 yuan. This is the most recommended option.
Subway: Take Subway Line 2 to Xuefu Avenue East Station (Exit 4), then transfer to Bus 864 to Qianhu Guest House Station; or take Subway Line 2 to Cuiyuan Road Station (Exit 4), then transfer to Bus 170 or 269 to Qianhu Guest House Station. The whole journey takes about 40 minutes and costs about 5 yuan.
Public Bus: Take High-speed Rail Bus Line 1 from the underground west bus stop at the north square to Qianhu Guest House Station or the campus of Jiangxi Science and Technology Normal University. The journey takes about 13–40 minutes and costs 5 yuan, with the last bus at 22:00.

➤ Departing from Nanchang Railway Station

Taxi/Online Car-hailing: The driving distance is about 11–14 kilometers, taking approximately 25–35 minutes, and the cost is about 30–50 yuan.
Subway: Take Subway Line 2 to Xuefu Avenue East Station (Exit 4), then transfer to Bus 864 to Qianhu Guest House Station; or take Subway Line 2 to Cuiyuan Road Station (Exit 4), then transfer to Bus 170 or 269 to Qianhu Guest House Station. The whole journey takes about 40 minutes and costs about 5 yuan.
Public Bus: Take High-speed Rail Bus Line 1 from the south bus stop at the west square of the railway station to the campus of Jiangxi Science and Technology Normal University or Qianhu Guest House Station. The journey takes about 34–60 minutes and costs 5 yuan, with the last bus at 22:00. Alternatively, take Bus 225 from Laofushan Garden Station to Qingyuan Road North Intersection Station or Nanchang University Qianhu Campus Station (about 1 hour and 20 minutes, 1 yuan). Note that the last Bus 225 departs at 19:30.

GUIDELINES

- **Onsite Oral Presentation**

1. The duration of a presentation slot is 15 minutes. Please target your lecture for a duration of about 10 minutes for the presentation plus about 5 minutes for questions from the audience.
2. Your punctual arrival and active involvement in each session will be highly appreciated.
   - Get your presentation PPT or PDF files prepared and backed up.
3. Laptops, projector & screen, laser sticks will be provided by the conference organizer.

- **Online Presentation**

1. Meeting Rooms

Zoom: 817 4009 8985

Password**: 052224**

Time Zone: China Standard Time (CST), UTC/GMT+8

2. Test Your Presentation

Date: **May 22, 2026**

Prior to the formal meeting, online presenters shall join the test room to ensure everything is on the right track. Please check your test Zoom Meeting ID on this program.

3. Oral Presentation

Please join the meeting room 10 minutes in advance.

Stay online during Keynote & Invited speeches and your own sessions. English Only during the conference. Rename your screen name before entering the room.

4. Follow us on We-chat

AGENDA OVERVIEW

All times in this schedule are listed in GMT+8,

May 22, 2026 | Friday		
Time	**Activity**	**Venue**
11:00-16:00	Onsite Sign-in & Conference Materials Collection	Lobby of Nanchang Qianhu Hotel
13:00-16:00	Online Sign-in & Equipment Testing	817 4009 8985 Password: 052224

May 23, 2026 | Saturday		
Time	**Activity**	**Venue**
	Opening Ceremony **Host: Cihui Yang,** Nanchang Hangkong University, China	
09:00-09:05	**Opening Remarks** **Jiyang Dai,** Nanchang Hangkong University, China	VIP Lounge 贵宾厅 3F
09:05-09:10	**Welcome Address** **William Wei Song,** Dalarna University, Sweden	VIP Lounge 贵宾厅 3F
09:10-09:20	**Group Photo**	
	Keynote Speech **Host: Sheng Zhang,** Nanchang Hangkong University, China	
09:20-10:00	**Keynote Speech** **Xiaoming Li**, Great Bay University, China ***Title: Big Data, Small Algorithm, and Little Mathematics-An Approach to Interdisciplinary Computing Education in the Age of AI***	VIP Lounge 贵宾厅 3F
10:00-10:20	**Coffee Break**	
10:20-11:00	**Keynote Speech** **Hong Rao,** Nanchang University, China ***Title: Deep Learning for Neuroimaging: Towards Accurate Auxiliary Diagnosis of Neurological Disorders***	VIP Lounge 贵宾厅 3F
11:00-11:30	**Keynote Speech** **Xiaohuan Wang,** Chinese Academy of Forestry, China ***Title: Smart Bamboo & Wood, Digital Construction: Big Data, Digital Twins and Digitalization of Wood and Bamboo Building Materials***	VIP Lounge 贵宾厅 3F

11:30-14:00	Lunch	Restaurant 1F
14:00-15:30	Technical Session 1 - Big Data Intelligence and Knowledge Discovery BT5010, BT5016, BT3003, BT6021, BT6018, BT5013	Meeting Room 306 3F
15:30-16:00	Coffee Break	
16:00-17:30	Technical Session 2: Smart Data Analytics and Applications BT6025, BT6028, BT6022, BT6029, BT6026, BT3002 - A	Meeting Room 306 3F
18:00	Dinner	Room 201 2F

May 24, 2026 | Sunday		
Time	Activity	Venue
10:00-12:00	Keynote Presentation Recording	817 4009 8985 Password: 052224
12:00-14:00	Break	
14:00-15:45	Online Session: Big Data Analytics and Scene Intelligence for Interdisciplinary Applications BT3001, BT4004, BT6027, BT5008, BT5009, BT5006, BT6024-A	817 4009 8985 Password: 052224

KEYNOTE SPEAKER

09:20-10:00 | VIP Lounge | 贵宾厅 3F

Xiaoming Li

Professor at Great Bay University, China
CCF Fellow, Ruisheng MOOC Chair Professor at Peking University

Biography: Xiaoming Li is a CCF Fellow, Ruisheng MOOC Chair Professor at Peking University, and currently Vice President (Second-level Head) and Chair Professor at Greater Bay University. He received his Ph.D. in Computer Science from Stevens Institute of Technology (USA) and B.Eng. from Harbin Institute of Technology. A pioneer in Web search, MOOCs, and educational technology in China, he led the development of the Tianwang Search Engine and founded the Chinese Web Archive (Web InfoMall). He served as Chair of Peking University's Department of Computer Science, Assistant President of Peking University, and Vice President of the China Computer Federation (CCF). His research spans web data mining, search engines, MOOCs, and computing education, with numerous awards including the CCF Wang Xuan Award, State Science and Technology Progress Award, and CCF Outstanding Education Award.

Speech Title: Big Data, Small Algorithm, and Little Mathematics-An Approach to Interdisciplinary Computing Education in the Age of AI

Abstract: I will start from my connections to big data, including a web search engine and a web archive, then proceed to an interdisciplinary undergraduate course involving social science and computing. I will show, as an educational approach, how data, model, and algorithms are orchestrated to understand and appreciate social sciences problems. In particular, I'll argue that the ability of orchestrating data, model, and algorithms is indispensable even in AI era, and must be trained rigorously.

KEYNOTE SPEAKER

10:20-11:00 | VIP Lounge | 贵宾厅 3F

Hong Rao

Professor at Nanchang University, China

Biography: Rao Hong, Ph.D. in Engineering, Second-Level Professor, Doctoral Supervisor, Young and Middle-Aged Key Teacher of Jiangxi Province, Teaching Master of Nanchang University, and currently Dean of the School of Software (School of Cybersecurity) at Nanchang University. She also serves as Vice Chairperson of Jiangxi Computer Federation and Jiangxi Digital Economy Society. She has long been engaged in artificial intelligence, data security technologies, and related fields.

Her main research interests include artificial intelligence, big data analysis and data mining, natural language processing, and data security. She has published over 100 research papers in top-tier international conferences and journals, including KDD, AAAI, and CVPR, and has edited 4 textbooks. In recent years, she has received one First Prize of Jiangxi Provincial Science and Technology Progress Award, and multiple Jiangxi Provincial Teaching Achievement Awards, including one Special Prize, one First Prize, and two Second Prizes.

Speech Title: Deep Learning for Neuroimaging: Towards Accurate Auxiliary Diagnosis of Neurological Disorders

Abstract: Neurological disorders such as brain tumors, Alzheimer's disease, and intracerebral hemorrhage pose significant challenges for early diagnosis and precise treatment due to their complex pathology and variable clinical presentations. Traditional imaging-based diagnosis relies heavily on radiologist experience and often suffers from subjectivity and low efficiency. This report presents a series of deep learning approaches developed by the research team at Nanchang University to address these challenges.

KEYNOTE SPEAKER

11:00-11:30 | VIP Lounge | 贵宾厅 3F

Xiaohuan Wang

Associate Professor at Chinese Academy of Forestry, China

Biography: She graduated from the Chinese Academy of Forestry in 2009 with a doctoral degree in Wood Science and Technology. She has long been engaged in research on the digital intelligent manufacturing of wood and bamboo structural materials, including the fabrication of wood-bamboo structural materials, prefabricated wall technologies for domestic timber, and hygrothermal performance research of wood and bamboo buildings.In recent years, she has focused on equipment development. She has carried out research and application of air-ground integrated bamboo harvesting and transportation technologies and equipment, continuous intelligent manufacturing production lines for laminated bamboo timber, machine vision and stress grading lines for plantation solid wood, and CNC machining centers for wood structure buildings.

Speech Title: Smart Bamboo & Wood, Digital Construction: Big Data, Digital Twins and Digitalization of Wood and Bamboo Building Materials

Abstract: Driven by global green low-carbon construction trends and digital upgrading of building materials, renewable, eco-friendly wood and bamboo building materials have gained prominent development momentum, especially after INBAR initiated the Bamboo for Plastic campaign. Nevertheless, the traditional industry suffers from extensive management, data isolation, insufficient performance regulation and incomplete full-lifecycle maintenance, making digital transformation imperative.This paper explores the integration pathways of big data and digital twin technologies, and builds a full-chain digital system covering raw material supply, smart production, performance inspection, construction and lifecycle management.With verified practical cases, it confirms the effects of relevant technologies in quality improvement, efficiency growth, low-carbon governance and industrial advancement. It also analyzes existing development bottlenecks and prospects the industry's green, digital, intelligent and international growth, offering references for industrial digital innovation and carbon neutrality implementation.

TECHNICAL SESSIONS

Session 1: Big Data Intelligence and Knowledge Discovery

Chair: **Yan Xu,** Jiangxi University of Finance and Economics, China & **Xiaohuan Wang,** Jiangxi University of Finance and Economics, China

14:00-15:30 | Meeting Room 306 3F

TAIK DETAILS	
Time	**Presentation**
14:00 - 14:15 BT5010	Title: DP-SIRS: A Privacy-Preserving Framework for Financial Risk Contagion Monitoring over Guarantee Networks Author(s): Ziang Wang, Jianping Zeng Presenter: Ziang Wang, Fudan University, China Abstract: Financial guarantee networks, where enterprises provide mutual guarantees for loans, are critical yet fragile structures susceptible to cascading risk contagion. Monitoring such contagion requires access to sensitive edge-level financial data (e.g., guarantee amounts), creating a fundamental tension between systemic risk oversight and commercial privacy. In this paper, we propose \textbf{DP-SIRS}, a novel collaborative framework that integrates the Susceptible-Infected-Recovered-Susceptible (SIRS) epidemiological model with differential privacy mechanisms for privacy-preserving financial risk contagion monitoring. Our framework introduces a center--node collaborative architecture: a central monitor observes network topology and node states but cannot access raw guarantee amounts, while enterprise nodes normalize and perturb their local edge weights via calibrated Laplace noise before uploading. To mitigate the utility loss induced by the necessary truncation of noisy values to a valid range, we further propose a truncation bias correction technique that analytically estimates and compensates for the systematic bias. We evaluate DP-SIRS on three synthetic network topologies---random graphs, small-world networks, and scale-free networks--under varying privacy budgets. Experimental results demonstrate that our bias-corrected mechanism achieves up to 35\% reduction in mean absolute error compared to the standard Laplace mechanism, while maintaining the same ε-differential privacy guarantee.
14:15 - 14:30 BT5016	Title: Prospective Analysis of Software Testing in Big Data Scenarios: From Traditional Testing to Agent-Based Testing Author(s): Jinchang Hu, Sen Yang, Feng Wu, Zhanwei Hui, Jun Wang, Chengying Mao Presenter: Feng Wu, Jiangxi University of Finance and Economics, China Abstract: The rapid development of big data, artificial intelligence, and distributed software systems is reshaping the context of software quality assurance. In big data scenarios, software systems are increasingly characterized by large-scale data processing, distributed execution, cross-platform collaboration, and dynamic runtime behavior, making conventional testing approaches centered on localized code verification increasingly inadequate. At the same time, recent advances in large

language models and agentic software engineering are creating new possibilities for more adaptive and intelligent testing. This paper reviews the evolution of software testing paradigms in big data scenarios, with particular attention to the transition from traditional testing to agent-based testing. It summarizes the background and motivation of testing in data-intensive environments, reviews related work on traditional software testing, intelligent unit testing, and agentic software engineering, and discusses how software testing has evolved from traditional and automated testing toward data-driven, intelligent, and agent-based paradigms. The paper further examines the application value of agent-based testing, including improved coverage, feedback integration, workflow adaptability, and process-level quality governance, while also discussing major challenges such as context understanding, long-horizon reasoning, trustworthiness, data quality, and integration with existing workflows. Possible future directions are also discussed, including observability-driven testing, multi-agent collaboration, unified engineering infrastructure, and governance mechanisms. Overall, this review highlights agent-based testing as an emerging direction in the broader transformation of software quality assurance under the joint influence of big data and artificial intelligence.

14:30 - 14:45
BT3003

Title: Data Weaving Methodology Based on Indicator Semantic Layer
Author(s): Nan Zhang, Yuxiao Zhao, Haosheng Lu, Mingli Chen, Ziqi Lu, Zhe Wu
Presenter: Nan Zhang, Nanjing Nari Digital service Co.,Ltd., China

Abstract: This paper addresses governance challenges arising from semantic conflicts, temporal fragmentation, and computational mismatches in cross-source heterogeneous data environments by proposing a metric-driven dynamic data fabric methodology. The approach establishes a three-layer semantic modeling framework for atomic, derived, and composite metrics, introducing three key innovations. First, we develop Indicator Description Logic (IDL), extending semantic web ontologies to overcome limitations in expressing computational rules while preserving ontological foundations. Second, we implement a Indicator Dependency Graph (MDG) enabling efficient lineage tracing and dynamic version control for large-scale relationships. Thirdly, based on the above research, we propose a conflict resolution mechanism that combines indicator semantics and relationship graphs to achieve high accuracy rates. This methodology provides a theoretical framework for data fabric that bridges business semantics with physical implementations through metric-centric weaving, ensuring semantic interpretability, decision verifiability, and path optimizability. Finally, some experimental results are conducted to verify the effectiveness and rationality of the method.

14:45 - 15:00
BT6021

Title: PoICRec: Attention-Based Interdisciplinary Collaborator Recommendation for Scientific Project Team Formation
Author(s): Yuansheng Zhong, Bohao Lin, Heqi Ouyang
Presenter: Heqi Ouyang, Jiangxi University of Finance and Economics, China

Abstract: Most existing studies on scientific collaborator recommendation focus on co-authorship prediction or general collaboration discovery, while the scenario of project team formation remains underexplored. To address this gap, this paper proposes a Project-oriented Interdisciplinary Collaborator Recommendation (PoICRec) model. PoICRec initially leverages a pre-trained language model to encode project titles and

keywords for capturing project-specific thematic requirements. It then constructs interdisciplinary features from scholars' historical project semantics and incorporates collaboration trust information to enhance candidate representations. Furthermore, a context-aware dynamic social-collaborative attention mechanism is introduced to differentiate the contributions of historical collaborators under the current project context. Finally, the model jointly optimizes ranking learning and trust-aware auxiliary supervision to generate Top-K collaborator recommendations for specific project scenarios. Experiments on completed project data from the National Natural Science Foundation of China (NSFC) show that PoICRec achieves competitive performance in project team formation, especially in cold-start settings and in identifying potentially valuable interdisciplinary candidates. The proposed method improves the modeling of the triadic relationship among project topic, principal investigator background, and candidate scholars, and provides practical support for project leaders seeking collaborators with topic relevance, disciplinary complementarity, and realistic collaboration feasibility.

15:00 - 15:15
BT6018

Title: Biomimetic Material Information Extraction and Knowledge Discovery Based on Large Language Models
Author(s): Mingyang Li, Junjie Wen, Zhun He, Zihan Wu, Yu Wu, Xiaoming Zhang
Presenter: Mingyang Li, Anhui University, China

Abstract: To address key challenges in biomimetic materials research-such as integrating multi-source heterogeneous data, correlating cross-scale information, and improving research efficiency-this study designs and implements an autonomous, end-to-end information extraction and knowledge discovery pipeline powered by Large Language Models (LLMs). Distinct from existing fragmented approaches that rely on decoupled tools or heavyweight infrastructures, our framework integrates intelligent keyword generation, a high-performance multi-source crawler cluster, and LLM-based semantic filtering into a unified workflow. A local Qwen-Plus model is deployed to automate the classification of 87,629 papers, significantly refining the corpus with high fidelity. Crucially, by migrating from dependency-heavy frameworks to the lightweight, graph-enhanced MiniRAG, we constructed a fully self-reliant question-answering system that balances domain-specific accuracy with computational efficiency. The resulting prototype validates the feasibility of driving intelligent upgrades in scientific research tools and provides robust, interpretable support for the design and application of biomimetic materials.

15:15 - 15:30
BT5013

Title: Skill Gap Analysis in Graduate Recruitment: A Weakly-Supervised Extraction Framework with CNN-MHA Hybrid Encoding
Author(s): Yuansheng Zhong, Yongyu Xie
Presenter: Yongyu Xie, Jiangxi University of Finance and Economics, China

Abstract: The accelerating advancement of artificial intelligence is reshaping the skill demands for graduates, yet existing studies predominantly rely on static or simplistic word-frequency methods, which fail to capture fine-grained skill dynamics. Moreover, the scarcity of manually annotated skill labels hinders large-scale recruitment data analysis. To address these limitations, this paper proposes a weakly supervised skill gap analysis framework. A BERT-CNN-BiLSTM-MHA-CRF model is constructed for fine-grained skill entity extraction. Using a corpus of 2.56 million entry-level job postings

from 2014 to 2024, the framework systematically identifies the evolution of skill demands and examines the divergent trajectories of hard and soft skills. The empirical results reveal significantly different depreciation rates between hard and soft skills: emerging competencies such as ‘large language models’ and ‘generative AI’ exhibit rapidly growing demand, while a number of traditional skills depreciate steeply. This study makes three principal contributions. First, weakly supervised learning is introduced into skill gap analysis, effectively addressing the label scarcity problem in large-scale recruitment data. Second, the decade-spanning temporal analysis provides empirical evidence for the asymmetric depreciation of hard versus soft skills, thereby enriching skill evolution theory. Third, the identification of emerging and declining skill demands offers data-driven insights for career planning, curriculum design, and recruitment strategy formulation.

TECHNICAL SESSIONS

Session 2: Smart Data Analytics and Applications

Chair: **Ziyue Tang**, Nanchang Hangkong University, China

16:00-17:30 | Meeting Room 306 3F

TAIK DETAILS	
Time	**Presentation**
16:00 - 16:15 **BT6025**	Title: A BERT-Based Framework for Sentence-Level Hallucination Detection Author(s): Wenxin Huang, Yanqing Liu, Wei Song Presenter: Wenxin Huang, Jiangxi University of Finance and Economics, China Abstract: With the widespread application of large-scale pre-trained language models in natural language generation tasks, the phenomenon of 'hallucination' in AI-generated content has increasingly become a critical issue affecting information credibility and decision-making quality. Large language models often suffer from hallucinations during generation-producing content unsupported by input contexts or external knowledge, which undermines text reliability. This paper addresses hallucinatory sentences in documents by proposing a BERT-based sentence-level hallucination detection method, which determines whether each segmented sentence contains 'unsubstantiated statements.' Methodologically, we employ BERT as a discriminative model, design an incremental fine-tuning mechanism and an interactive feedback module, and evaluate the model's detection performance under limited annotated data. Experimental results demonstrate that the BERT model exhibits satisfactory semantic modeling capability for hallucination detection; however, the model generally suffers from low confidence and still faces difficulties in identifying hallucination types that are heavily dependent on contextual information. This paper provides an operational formal definition, a systematic experimental framework, and an extensible evaluation system for the hallucination detection task.
16:15 - 16:30 **BT6028**	Title: A BERT Based LLM System for Legal Cases with Feedback Mechanism Author(s): Shiyu Du, Yan Xu, Wei Song Presenter: Shiyu Du, Jiangxi University of Finance and Economics, China Abstract: To address limited efficiency, weak boundary-case recognition, and insufficient legal semantic understanding in legal case assistance, this paper proposes a BERT-based dual-task experimental framework for two representative legal scenarios: civil defamation text classification and criminal guilty/not guilty prediction. Methodologically, BERT-Base-Chinese is adopted as the base model, and a human-in-the-loop interactive feedback mechanism is introduced on top of standard fine-tuning to support iterative optimization through human evaluation, immediate correction, and parameter updating. Experimental results show that pre-trained language models can establish a basically usable classification capability on typical legal text samples, but they remain vulnerable to surface lexical cues and high-frequency features in boundary cases such as sentiment traps, sarcastic or insinuating expressions, justifiable

defense, injury-threshold scenarios, and accidental incidents. Further results demonstrate that human feedback can effectively correct erroneous decision boundaries and improve the model's ability to identify complex legal semantics and rule conditions. These findings provide preliminary evidence that pre-trained language models can support exploratory legal text classification, but their practical effectiveness still depends on task design, boundary-sample construction, dataset realism, and human-machine feedback mechanisms.

16:30 - 16:45
BT6022

Title: Table Structure Recognition: A Robust Algorithm for Document Table Structure Reconstruction
Author(s): Desheng Yang, Cihui Yang
Presenter: Desheng Yang, Nanchang Hangkong University, China

Abstract: Table structure recognition is a critical task in document analysis. This paper presents a robust algorithm that processes table structures through distance threshold calculation, cell sorting and row grouping, reference row-based bidirectional diffusion alignment, and intelligent cell segmentation. The algorithm employs a reference row identification strategy with upward and downward diffusion for cross-row alignment, effectively handling merged cells and irregular layouts. Experimental results demonstrate the algorithm's effectiveness in accurately reconstructing table structures from various document formats.

16:45 - 17:00
BT6029

Title: Wood Moisture Content Prediction Method Using Multi-Sensor Spatial Feature Fusion
Author(s): Tianhao Liu, Xiaohuan Wang, Yan Xu, Wei Song*
Presenter: Tianhao Liu, Jiangxi University of Finance and Economics, China

Abstract: Moisture content (MC) is critical for timber structure durability and safety. Conventional single-point monitoring cannot capture spatial microclimate heterogeneity, limiting prediction accuracy. This study proposes a multi-sensor spatial feature fusion framework with vertical-gradient and horizontal-difference features integrated into a Group Method of Data Handling (GMDH) model. Experiments on 8,927 records show 29.5% MAE reduction, offering a practical solution for smart timber building moisture early warning.

17:00 - 17:15
BT6026

Title: A Hybrid CNN-LSTM-RF Model for Predicting Moisture Content of Reconstituted Bamboo Building Materials
Author(s): Yan Xu, Wenjie Xu, Xiaohuan Wang, Wei Song
Presenter: Wenjie Xu, Jiangxi University of Finance and Economics, China

Abstract: Moisture content dominates the structural stability and service life of reconstituted bamboo building materials. Traditional methods fail to capture complex nonlinearity and temporal dependence, resulting in low prediction accuracy. This paper proposes a hybrid CNN-LSTM-RF model, combining spatial feature extraction, temporal dependency learning and robust regression. Validated on real monitoring data, the model outperforms mainstream baseline approaches with superior precision and generalization. It provides an effective solution for intelligent monitoring and safety evaluation of reconstituted bamboo structures.

17:15 - 17:30
BT3002 - A

Title: Research on Entity Alignment Method in Comparison Opinion Mining Based on Multi-Feature Fusion
Author(s): Na Wei
Presenter: Na Wei, Chang'an University, China

Abstract: Comparison opinion mining from user perspective is significant for gaining deep insights of strengths and weaknesses of a product relative to its competitors, clarifying market positioning of a product or a brand, and formulating product redesign and innovation strategies. However, due to the issue of colloquialism in the opinions expressed by social media users, there is a lack of standardization in the entity names in the comparison opinion text. After mining comparison opinions, implementing entity alignment is an important guarantee for achieving accurate global comparison of products. Existing traditional entity alignment methods and knowledge-based entity alignment methods overly rely on manual feature engineering and cannot capture deep semantic information, making it impossible to align sparse entities. This study proposes an entity alignment method based on multi-feature fusion. We first generate entity direct semantic embeddings based on the Transformer model and identifies high confidence aligned entity pairs based on entity embeddings. Then, we generate a comparison network based on the mined comparison tuples, and integrate features including the structure of the comparative network graph, comparative attributes, and entity semantic information, and implement unknown entity alignment based on the graph attention network model and Gromov-Wasserstein Distance. Finally, the effectiveness of the model was validated based on two real-world datasets, and the results showed that the proposed method in this study had a significant improvement compared to baseline methods.

ONLINE SESSIONS

Online Session: Big Data Analytics and Scene Intelligence for Interdisciplinary Applications

Chair: **Hui Zhang**, Southwest University of Science and Technology, China

14:00-15:30 | 817 4009 8985 Password: 052224

TAIK DETAILS	
Time	**Presentation**
14:00 - 14:15 **BT3001**	Title: Construction and Application of Visual Narrative Scene Driven by Big Data in the Field of Education Author(s): Cong Wang, Zhaoying Chen Presenter: Cong Wang, University of Jinan, China Abstract: This article focuses on the construction and application of visual narrative scenes in the field of education driven by big data. It elaborates on the core concepts of big data and visual narrative and their value in education, analyzes the key technologies for constructing visual narrative scenes, explores their application modes in different educational scenarios. The aim is to provide theoretical support and practical reference for the digital and information-based transformation of education.
14:15 - 14:30 **BT4004**	Title: Application of Big Data in the Evaluation of Mathematical Quotient Author(s): CHENGLONG LI, ZHAOYING CHEN, ANXIN ZHU Presenter: CHENGLONG LI, University of Jinan, China Abstract: Mathematical Adversity Quotient is the core ability of individuals to cope with adversities in mathematics learning. Leveraging the application advantages of big data technology in the field of mathematics education-specifically by analyzing massive behavioral trajectory data generated during students 'mathematical learning to assist in questionnaire design and result verification-this study explores its structural dimensions using principal component factor analysis. Taking high school students as research subjects, an empirical analysis was conducted through the compilation of the Mathematical Adversity Quotient Questionnaire. The results show that several common factors were extracted from the questionnaire through principal component factor analysis, with the cumulative variance explanation rate reaching an acceptable level; after factor rotation, all item loadings were appropriate with no cross-dimensional loadings, indicating good reliability and validity of the scale. The constructed evaluation model of Mathematical Adversity Quotient provides tool support for the scientific assessment of adversity quotient levels in educational practice and improves the non-intellectual factor evaluation system.

Title: Multifactor Sensitivity Analysis of Timber Moisture Content in a Small Wooden Cabin
Author(s): Yan Xu, Zhen Chen, Xiaohuan Wang, Wei Song
Presenter: Zhen Chen, Jiangxi University of Finance and Economics, China

14:30 - 14:45
BT6027

Abstract: Wood moisture content (MC) is critical to timber structure durability. To quantify the coupled effects of hygrothermal factors, this paper proposes a BO-XGBoost model for MC prediction using temperature, humidity, absolute humidity and dew point temperature as inputs. Three sensitivity analysis methods (DSA, Sobol, Morris) are adopted to identify factor influence weights. Results show that BO-XGBoost achieves high prediction accuracy, and temperature & relative humidity are the dominant factors. This work provides a quantitative basis for timber structure moisture-proof design and monitoring.

Title: Boundary-Aware Object Refinement and Mixed-Only Graph-Regularized FCLS for Multispectral Subpixel Mapping
Author(s): Jiang Lu, HangYu Liu, ChiCheng Ma, JiaQi Liu, ChengZhi Tang, YaXing Sun
Presenter: Jiang Lu, Nanjing University of Science and Technology, China

14:45 - 15:00
BT5008

Abstract: Abstract—Mixed pixels in multispectral images blur boundaries, suppress small targets, and distort class proportions, which limits subpixel mapping. Existing object-based methods such as UO-SPM improve abundance estimation by using object segmentation and secondary unmixing, but still rely on fixed thresholds and fixed erosion for pure/mixed identification. We propose BA-MF, which combines boundary-aware object refinement with mixed-only prior-constrained unmixing. The method refines object priors from Landsat spectra, initial soft classification, and boundary cues, identifies pure and mixed pixels adaptively, and re-estimates only mixed pixels by constrained FCLS with prior consistency and spatial smoothing. The refined priors are further used by PSA, MRF, RBF, and HNN. Experiments show improved full-region unmixing, mixed-region recovery, and final mapping accuracy over UO-SPM, with more stable boundary preservation and better small-target reconstruction. Index Terms-subpixel mapping, object refinement, boundary awareness, mixed pixels, FCLS, spectral unmixing

Title: Research on Integrating Mathematical Culture into Personalized Teaching Strategies Empowered by Big Data Technology
Author(s): Jing Liang, Zhaoying Chen
Presenter: Jing Liang, University of Jinan, China

15:00 - 15:15
BT5009

Abstract: Based on the requirements of the new era's mathematics teaching reform and core literacy cultivation, and in light of the practical needs of teaching students in accordance with their aptitudes in mathematics teaching, this paper explores the implementation strategies of integrating mathematics culture into personalized teaching under the support of big data. The article analyzes the multiple values of integrating mathematical culture into personalized teaching, and at the same time reveals the current practical predicaments in mathematics teaching, such as the lack of data support for the personalized design of cultural infiltration and the disconnection between the application of technology and teaching integration. On this basis, the internal mechanism of the personalized infiltration of mathematical culture driven by

big data is expounded, and the key role of big data in personalized teaching, providing personalized learning resources and dynamic tracking of students' learning process is clarified. Finally, from three dimensions: teaching content reconstruction, teaching method innovation, and teacher ability development, a practical path for integrating mathematical culture into personalized teaching supported by big data is systematically constructed. The aim is to achieve the empowerment of big data for integrating mathematical culture into personalized teaching, thereby effectively promoting the personalized development of students' mathematical cultural literacy.

15:15 - 15:30
BT5006

Title: CGLANet: A Context-Aware Global-Local Self-Attention Network for Efficient Remote Sensing Image Scene Classification
Author(s): LONGXI XU, MEIHUI XU, HAO YUAN
Presenter: LONGXI XU, University of Jinan, China

Abstract: In recent years, remote sensing image classification has gained significant attention due to its wide-ranging applications in environmental monitoring, urban planning, and natural resource management. Traditional Convolutional Neural Networks (CNNs) have limitations in capturing global context information, while Vision Transformers (ViT) excel in global feature extraction via self-attention but suffer from high computational complexity and redundant parameters. To address these issues, we propose CGLANet, a Context-Aware Global-Local Self-Attention Network that integrates the strengths of CNNs and Transformers. CGLANet utilizes hierarchical convolutions for multi-scale feature extraction and combines global and local self-attention mechanisms to enhance both global context modeling and local feature extraction. The model significantly reduces computational complexity and parameter redundancy, ensuring efficient processing without compromising classification accuracy. Experimental results on the AID, UCM, and NWPU datasets show that CGLANet outperforms state-of-the-art methods in terms of classification accuracy, computational efficiency, and model generalization. Specifically, CGLANet achieves 95.83% accuracy on the AID dataset with 20% training data, 98.40% on the UCM dataset with 50% training data, and 92.75% on the NWPU dataset with 10% training data. These results confirm CGLANet as a robust and efficient solution for remote sensing image scene classification.

15:30 - 15:45
BT6024-A

Title: Exploring Energy Efficiency and Indoor Air Quality in Residential Buildings through Data Analysis of Ventilation Systems: A Case Study on Dalarnas Villa
Author(s): Qinming Lin, Lianzhicong Zhang, Yurong Zhu, Xiaohuan Wang, William Wei Song
Presenter: Qinming Lin, Shanghai University, China

Abstract: This paper investigates energy efficiency and indoor air quality (IAQ) in residential buildings by analysing ventilation systems. The primary goal is to develop an optimized flowrate schedule that balances CO2 concentration control and energy consumption. Taking the Dalarnas Villa as a case study, the research employs data-driven and physical modelling approaches to predict indoor CO2 levels and optimize ventilation schedules. Time series forecasting models, including LSTM, LSTM Attention, and XGBoost, were developed to predict future CO2 concentrations based

on historical data. Among them, the LSTM+Attention model demonstrated the best predictive performance. Furthermore, a genetic algorithm (GA) was employed to generate the Optimized Schedule based on predicted CO2 levels, achieving significant energy savings while maintaining acceptable IAQ. The Optimized Schedule was evaluated through scenario-based analysis, demonstrating improved energy efficiency compared to traditional methods. This study demonstrates the feasibility of combining machine learning-based forecasting with evolutionary optimization to proactively control residential ventilation systems in the studied case, while transfer to other rooms or buildings requires site-specific recalibration.

MEMO