

2024 9th International Conference on Big Data Analytics (ICBDA 2024)

**Tokyo, Japan
16-18 March 2024**



**IEEE Catalog Number: CFP24TPE-POD
ISBN: 979-8-3503-5297-9**

**Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP24TPE-POD
ISBN (Print-On-Demand):	979-8-3503-5297-9
ISBN (Online):	979-8-3503-5296-2
ISSN:	2832-3726

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2024 the 9th International Conference on Big Data Analytics (ICBDA 2024)

Table of Contents

Preface	viii
Conference Committees	ix

✂ Machine Learning Models and Algorithms

Predicting COVID-19 Transmission in Southern California with Machine Learning Methods	1
<i>Han Li, Ran Wei, Wenyu Wang, Nanpeng Yu</i>	
Integrating Latent Dirichlet Allocation and Gradient Boosting Tree Methodology for Insurance Product Development Recommendation	11
<i>Wei-Yu Chen, Tzu-Chien Wang, Ruey-Shan Guo, Chialin Chen</i>	
An AI Framework for Predicting the Winner of the Grammys	20
<i>Rushabh Musthyala, Abhishek Narayanan, Anirudh Nistala, Anasse Bari</i>	
Dual-GRU-GNN for Dynamic Blockchain Network Sharding	26
<i>Xianghe Ye, Qingxu Di, Xinyuan Du, Ayong Bao, Shusong Xing</i>	
Integrating Spatio-Temporal Information for Forecasting Financial Indicators of Listed Companies	34
<i>Yijie Teng, Rongmei Yang, Linyuan Lü</i>	

✂ Data Based Intelligent Application Platform and Information Mssanagement

Visualization Study on Suitability Evaluation of Land Spatial Agricultural Production in the Yangtze River Delta Based on Remote Sensing Big Data	41
<i>Duanqiang Zhai, Jian Zhuo, Yanyun Mao</i>	
Exploration and Implementation of Data Processing Methods in Admission and Enrollment for Adult Higher Education Institutions	46
<i>Xuping Lin, Bin Wang, Yuejiang Sun, Lei Li</i>	
Enhancing Accuracy in LLM for Estimating Human Intentions: A Data-Driven Self-Validation Approach	51
<i>Yudai Yamazaki, Daisuke Kamisaka</i>	

Application of Data Anonymization to Smartphone Apps Usage Logs 60
Naoya Tabata, Maiko Shigeno

Towards Understanding Low Energy Nuclear Reactions: Structuring the Literature and Applying Data Analytics
..... 70
*Anasse Bari, Sneha Singh, Yvonne Wu, Tanya Pushkin Garg, Emos Ker, Charles Wang, Benjamin Kang,
Adelina Simpson, Dongjoo Lee, David Nagel*

Graph Similarity-Based Data Management Platform for Advancing Material Informatics..... 85
Amit Kumar, Hideya Yoshiuchi

✂ AI based Image Processing Technology and Methods

Text-Guided Co-Modulated Generative Adversarial Network for Image Inpainting 92
Chunjiang Fu, Minghao Wang, Qinghao Hu, Liang Zhao

Construction and Analysis of Knowledge Graph Based on Industrial Fault Diagnosis Data 98
Yanzhi Zhang, Weiqiang Qu, Zeyang Chen, Jiayi Zhang, Boyuan Zheng, Shenghui Shi

Ship Navigational Status Classification based on the Geometrical and Spatiotemporal Features of the
AIS-generated Trajectory 103
Wayan Mahardhika Wijaya, Yasuhiro Nakamura

A Synergistic Large Language Model and Supervised Learning Approach to Zero-Shot and Continual Activity
Recognition in Smart Homes 113
Naoto Takeda, Roberto Legaspi, Yasutaka Nishimura, Kazushi Ikeda, Thomas Plötz, Sonia Chernova

Vibration Data Mining and Machine Learning for Anomaly Detection of Cycling Trails Using Instrumented Bike
..... 123
Chun-Hsing Ho, Kewei Ren

Pseudo-Label Guided Partially View-aligned Clustering 128
Songtao Wu, Ruixin Ma, Qiongjie Xie, Liang Zhao

Learning M-Order Spectrum Graphs to Identify Unknown Chemical Compounds from Infrared Spectroscopy
Data 134
Gyoung S. Na, Ye Cheol Rho

✂ Software and Information Technology

A Corporate Financial Statement Fraud Model Based on Extracting Correlated Topic Features 144
Ching-Hsue Cheng

Typhoon-DIG: Distinguishing, Identifying and Geotagging Typhoon-Related Social Media Posts in Taiwan...	149
<i>Ben-Xiang Li, Chih-Yuan Chen</i>	
Mining Organization-Related Fine-Grained Information: A Bias-corrected Multi-Turn Question Answering Based Approach	157
<i>Minjie Yuan, Yajun Ren, Haoda Qian, Xue Mao, Xina Zhang, Qiudan Li</i>	
Risk Sensitive Distributional Soft Actor Critic for Portfolio Management.....	163
<i>Jun Liu, Yanlong Wang, Shao-Lun Huang, Danny D. Sun</i>	
Adaptive Cognitive Diagnosis under Distribution Shift with a Double-level Adversarial Framework.....	168
<i>Yuhong Zhang, Chenyang Si</i>	
The Dynamic Routing Load Balancing Algorithm Based on Blockchain Platform	174
<i>Yifei Luo, Xi Zhang, Xiaofeng Jia</i>	
Sentiment Analysis and Emotion Detection of COVID-19 Geo-Tagged Twitter Data.....	180
<i>Sujing Wang, Sambeg Dhakal, Bhuwan Prasad Upadhyay</i>	
Thermal Comfort in Air Conditioning Based on Bayesian Inference Using the EAP Estimation	186
<i>Yohei Saika, Masahiro Nakagawa</i>	
A Novel Differential Privacy Protection Model for Publishing Time Series of Classification Task	191
<i>Ai Ran, Duo Li, Xinyue Ren, Jinnan He, Fengchi Yuan, Wai Kin (Victor) Chan</i>	
∅ Data Model and Database	
Research and Data Analysis on Relationship Extraction Methods Based on Multi-Domain Texts	197
<i>Shenghui Shi, Zhangyan Li, Lin Yin, Zhaoying Chai, Kechen Fan, Daguang Jiang, Liangfei Zheng, Dongqi Huang</i>	
Universal Data Model as a Way to Build Multi-paradigm Data Lakes.....	203
<i>Artem A. Sukhobokov, Yury E. Gapanyuk, Artem A. Vetoshkin, Alexandra R. Mironova, Daniil R. Nikolskiy, Maria A. Morozevich, Nikita A. Klyukin, Rodion A. Afanasev, Dmitriy S. Lakhvich</i>	
Driving Biotech Innovation: An analysis of RD&I models and Big Data Strategies in Open Innovation Frameworks.....	212
<i>Jayson Luis da Silva Ribeiro, Anapatricia Morales Vilha, Enda Fallon, John Mee, Thiago Braga Rodrigues</i>	
Techniques used in time series databases and their internals	218
<i>P K Navin Shrinivas, Mohamed Ayaan, Anirudh Rowjee, Vishal CE, Animesh Giri</i>	
MySQL-Pot: A LLM-Based Honey-pot for MySQL Threat Protection	227
<i>Yuqi Hu, Siyu Cheng, Yuanyi Ma, Shuangwu Chen, Fengrui Xiao, Quan Zheng</i>	

∅ Neural Network Theory and Models in Digital Images

Connectivity-aware Experience Replay for Graph Convolution Network-based Collaborative Filtering in Incremental Setting233

Xin Fan, Fan Mo, Chongxian Chen, Changhao Bai, Hayato Yamana

Predicting Orthopedic Surgery Types Using Convolutional Neural Networks.....243

Li-Ling Wang, Chin-Hao Hsu, Tan-Yun Tsai, Li Huang, Po-Chou Shih

Use of Deep Neural Networks for Uncertain Stress Functions with Extensions to Impact Mechanics248

Garrett Blum, Ryan Doris, Diego Klabjan, Horacio Espinosa, Ron Szalkowski

Feature Extraction of Gait Data via PCA and Personal Identification Using Neural Networks258

Kazuki Nishimura, Akio Ishida, Jun Murakami, Naoki Yamamoto

∅ Blockchain Technology and Security in Modern Integrated Information Systems

Web3.0 Generative Art- A Co-creation System in Environmental Sustainability Concept Form264

Chiung-Hui Chen

DBNode: A Decentralized Storage System for Big Data Storage in Consortium Blockchains.....270

Narges Dadkhah, Xuyang Ma, Katinka Wolter, Gerhard Wunder

Attributes of a Blockchain-based Halal Tracking and Tracing System: A Multi-Criteria Decision-Making Approach280

Yandra Rahadian Perdana, Eko Heru Prasetyo

Challenges and Opportunities of Integrating Non-Fungible Tokens (NFTs) and Self-Sovereign AI (SSAI) in Blockchain-Based Metaverse Projects288

Christopher G. Harris

Anomaly Detection Method for Small-Scale Blockchain Transaction.....297

Shin Morishima

∅ Big Data Model, Data Analysis and Application

Adaptive Shapley: Using Explainable AI with Large Datasets to Quantify the Impact of Arbitrary Error Sources305

Birk Martin Magnussen, Maik Jessulat, Claudius Stern, Bernhard Sick

Mapping Big Data Analytics and Supply Chain Resilience Research Nexus: A Bibliometric Study311

Aziz Hmioui, Yasser El Ouarrak

Exploring a Big Data-Driven Dialogue Generation Method for Disease Queries using RASA.....316
Xinyue Quan, Xiaohan Dai, Xiang Xie

Implementing an Integrated Assistance System for Stock Investment with Big Data Environments.....322
Yu-Chieh Cheng, Yun-Hsuan Chang, Yu-Chi Lin, Hsin-Jou Shih, Chung Yung

Scalable Analysis of English Dictionary Files on HPC Systems Big Data Platform.....328
Jayanth C, Adarsh U, David De Hilster, Hugo Watanuki, Shobha G, Jyoti Shetty

Enhancing Open-Source Intelligence: Introducing the Functional Intelligence Tool (FINT).....334
Laurens H.F. Müter, Jacco van der Plaat, Anna Zoet, Christof van Nimwegen, Remco C. Veltkamp

Towards a Comprehensive Framework for Big Data in Higher Education344
Taher Omran Ahmed

Big Data, Spatial Economical Model and Ranking Size Rule352
Li Zhang

High Fluctuation Based Recursive Segmentation for Big Data.....358
Desmond Fomo, Aki-Hiro Sato