

2025 IEEE 12th International Conference on Data Science and Advanced Analytics (DSAA 2025)

**Birmingham, United Kingdom
9-12 October 2025**



**IEEE Catalog Number: CFP25DSB-POD
ISBN: 979-8-3315-1180-7**

**Copyright © 2025 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:	CFP25DSB-POD
ISBN (Print-On-Demand):	979-8-3315-1180-7
ISBN (Online):	979-8-3315-1179-1
ISSN:	2472-1573

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

TABLE OF CONTENTS

DyCAF-Net: Dynamic Class-Aware Fusion Network.....	1
<i>M. Abrar Jahin, Shahriar Soudeep, M. F. Mridha, Nafiz Fahad, M. Jakir Hossen</i>	
Drone Swarm Sensitivity Estimation using Bayesian Theory for Search and Rescue Operations	11
<i>Luciano Lima, Ali Ghaderi, Fabio Andrade, Carlos Pfeiffer, Youcef Djenouri, Marcos Moura</i>	
Data Driven Optimum Cyberattack Mitigation	21
<i>Erol Gelenbe, Mohammed Nasereddin</i>	
Fast Training on Dynamic Heterogeneous Information Network for Fake News Detection.....	30
<i>Jinho Go, Aldhytha K. Sari, Jiaojiao Jiang, Shuiqiao Yang, Sanjay K. Jha</i>	
Auto-CEn: AutoML for Classifier Ensembles - Diversity-Based Classifier Selection and Decision Fusion Optimization	40
<i>Julius Voggesberger, Peter Reimann, Dennis Treder-Tschechlov, Bernhard Mitschang</i>	
Expert Detection on Crowdsourcing Forums via Sheaf Laplacian.....	50
<i>Mehmet E. Aktas, Iraj Moradi, Ibrahim Cosar, Yateeka Goyal, Asya Alyaz, Esra Akbas, Mehmet Ahsen</i>	
A Survey on Current Trends and Recent Advances in Text Anonymization	60
<i>Tobias Deußner, Lorenz Sparrenberg, Armin Berger, Max Hahnbüch, Christian Bauckhage, Rafet Sifa</i>	
From Engineering Diagrams to Graphs: Digitizing P&IDs with Transformers	69
<i>Jan M. Stürmer, Marius Graumann, Tobias Koch</i>	
Evaluating Anomaly Detection Algorithms: The Role of Hyperparameters and Standardized Benchmarks	80
<i>Simon Klüttermann, Shubham Gupta, Emmanuel Müller</i>	
Unveiling Fairness and Performance of Causal Discovery	88
<i>Sónia Teixeira, Ana R. Nogueira, João Gama</i>	
Cybersecurity Recommendations for Planning and Securing Port 4.0 and Maritime Industry Against Cyberattacks.....	98
<i>Nitesh Bharot, Priyanka Verma, Rutvij H. Jhaveri, John G. Breslin</i>	
HyperKD: Distilling Cross-Spectral Knowledge in Masked Autoencoders via Inverse Domain Shift with Spatial-Aware Masking and Specialized Loss.....	108
<i>Abdul Matin, Tanjim B. Faruk, Shrideep Pallickara, Sangmi L. Pallickara</i>	
The Value of Diversity for Dealing with Concept Drift in Class-Imbalanced Data Streams.....	119
<i>Chun W. Chiu, Leandro L. Minku</i>	
Multi Algorithm Selection and Hyperparameter Optimization for Automated Clustering	129
<i>Georg S. Schlake, Max Pernklau, Christian Beecks</i>	
Small and Fast LLMs on Commodity Hardware: Post-Training Quantization in llama.cpp.....	138
<i>Lorenz Sparrenberg, Tobias Deußner, Armin Berger, Rafet Sifa</i>	
Evaluating Defensive Influence in Multi-Agent Systems using Graph Attention Networks.....	148
<i>Gregory Everett, Ryan J. Beal, Tim Matthews, Timothy J. Norman, Sarvapali D. Ramchurn</i>	

Abstention is All You Need	158
<i>Erik Schönwälder, Christian Falkenberg, Claudio Hartmann, Wolfgang Lehner</i>	
Leveraging Causal Discovery to Tackle Complexity in Model-Based Anomaly Detection: Case-Study from Blast Furnace Operation	168
<i>Matej Vukovic, Belgin Mutlu, Thomas Kristan, Petra Krahwinkel, Christian Tauber, Stefan Thalmann</i>	
Reasoning LLMs in the Medical Domain: A Literature Survey	178
<i>Armin Berger, Sarthak Khanna, Lorenz Sparrenberg, Tobias Deußner, David Berghaus, Rafet Sifa</i>	
Semi-Supervised Relation Extraction Informed by Area Under the Margin Ranking and Large Language Models	188
<i>Nikita Gautam, Bipin Paudel, Doina Caragea, Cornelia Caragea</i>	
An Event Study Framework for Analyzing Bidirectional Causality Between House Prices and Premium Café Entry	198
<i>Sachin Date</i>	
Hyperon: An Online Hyperparameter Tuning Approach for Data Stream Learning	208
<i>Sadia Tabassum, Leandro L. Minku</i>	
Diverse and Plausible Counterfactual Explanations for Time Series via Latent Space Optimization	218
<i>Peiyu Li, Omar Bahri, Pouya Hosseinzadeh, Soukaina F. Boubrahimi, Shah M. Hamdi</i>	
First Investigation of Deep Learning for Intraoperative Gauze Segmentation in Minimally Invasive Abdominal Surgery	228
<i>Priya Tomar, Maximilian Broß, Philipp Feodorovici, Jan Arensmeyer, Philipp Leifels, Aditya Parikh, Hanno Matthaei, Christian Bauckhage, Helen Schneider, Rafet Sifa</i>	
Reasoning with Knowledge Graphs for Trustworthy Course Recommendation	239
<i>M. A. Z. Khan, Dongsheng Luo, Agoritsa Polyzou</i>	
Scalable Deep Subspace Clustering Network.....	249
<i>Nairouz Mrabah, Mohamed Bouguessa, Sihem Sami</i>	
CA-HIL: A Competency-Aware Framework for Bridging Automation and Human Expertise in Data Science Projects	259
<i>Maike Holtkemper, Christian Beecks</i>	
RST-NDAS:RST Nuclei for Discourse-Aware Abstractive Summarization.....	268
<i>Mohammadreza Sediqin, Shlomo E. Argamon</i>	
Data Subcard: Evaluating Privacy, Fairness, Quality, and Protection in Tabular Data, as Part of the System Cards Framework.....	273
<i>Tadesse K. Bahiru, Carlos Ordonez, Ioannis A. Kakadiaris</i>	
Towards LLM-Guided Healthcare Dataset Harmonization	279
<i>Christos Smailis, Carlos Ordonez, Ioannis A. Kakadiaris</i>	
Improving the End to End Purchasing Process in Indirect Procurement	283
<i>Rama Reddy, Ravi Ilkal, Snehal Soni</i>	
Generalization in Audio Deepfake Detection: Evaluating ASR Encoder-Based Feature Extraction.....	286
<i>Karla Schäfer, Martin Steinebach</i>	

Mind the Gap: Validating System Requirements for Competence-Based Decision Support in Data Science Projects.....	291
<i>Maike Holtkemper, Christian Beecks</i>	
REFINE: Enabling Efficient and Trustworthy Modeling of Financial Networks via GNN-to-MLP Knowledge Distillation.....	296
<i>Vipul K. Singh, Jyotismita Barman, Sandeep Kumar, Jayadeva</i>	
Towards an Explainable-AI Enhanced Predictive Analytics System for Collaborative Learning in Smart Educational Environments	301
<i>Rawia Bdiwi</i>	
A New Framework for Causal Inference Without Missing Data Analysis: Bayes Optimal Treatment Effect Estimation	306
<i>Kohei Horinouchi, Shimpei Onishi, Toshiyasu Matsushima</i>	
Spatially Adaptive PM _{2.5} Estimation in Low-Sensor Regions using Variational Gaussian Processes.....	311
<i>Shrey Gupta, Avani Wildani, Yang Liu</i>	
HyperNeRFGAN: Camera-Free 3D Scene Generation via Hypernetwork-Driven Neural Radiance Fields	316
<i>Adam Kania, Artur Kasymov, Jakub Kosciukiewicz, Artur Górak, Marcin Mazur, Maciej Zieba, Przemyslaw Spurek</i>	
Reproducibility and Case Sensitivity of LLMs for Anonymizing Depressed Tweets	325
<i>Sameen Mansha, Hamza Mahmood, Anne Håkansson, Faisal Kamiran, Vladimir Vlassov</i>	
Detection (Reconstruction) of Long-Lived Particle Tracks in LHCb Experiment, CERN.....	335
<i>Piotr Kulczycki, Grzegorz Golaszewski, Tomasz Szumlak, Szymon Lukasik</i>	
PSM: Proactive Spill Mitigation in PySpark	345
<i>Amin Karami</i>	
Explainability of Molecular Graph Neural Network	355
<i>Ataollah Kamal, Matej Hladiš, Jérémie Topin, Marc Plantevit, Sébastien Fiorucci, Céline Robardet</i>	
Anomaly Detection Under Contaminated Data: A Weighted Iterative Refinement Framework for Health Monitoring	365
<i>Stefano Donné, Jesse Davis, Filip Van Utterbeeck, Mathias Verbeke</i>	
Atrous-CNN with Hierarchical-Based Training Strategy Approach for Decentralized Tasks	376
<i>Antoni Jaszcz, Agnieszka Polowczyk, Alicja Polowczyk, Katarzyna Wiltos, Dawid Polap, Marcin Wozniak</i>	
Student-Perceived Cognitive Load of LLM-Generated Programming Exercises.....	384
<i>Muhammad F. A. Khan, Ludia E. Feri, Ha Nguyen, Hamid Karimi</i>	
CACTUS: Cross-Aligned Counterfactual Explanation for Time Series Classification.....	396
<i>Pouya Hosseinzadeh, Peiyu Li, Omar Bahri, Soukaina F. Boubrahimi, Shah M. Hamdi</i>	
A Comparative Study of Non-Linear Modelling Capabilities of T-S Fuzzy Models and Back-Propagation Neural Networks	406
<i>Linxiang Li, Kainan Liu, Xiaojun Ban, Shengkun Xie</i>	
3D Point Cloud Classification using Graph Convolutional Network with Multi-Scale Pooling.....	416
<i>Alicja Polowczyk, Agnieszka Polowczyk, Antoni Jaszcz, Marcin Wozniak, Dawid Polap</i>	

Offshore Asset Inspection Redefined: Expert- Validated Deep Learning for Critical Defects	424
<i>Urmila Gurung, M. S. Mekala, Carlos F. Moreno-Garcia, Nikolaus Zolnhofer, Barry Marshall, Elyan Eyad</i>	
ADT++: Advanced Adversarial Distributional Training with Class-Wise Robustness	434
<i>Samer Y. Khamaiseh, Deirdre Jost, Anas Alsobeh, Abdullah Al-Alaj, Honglu Jiang</i>	
Enhancing Touchless Sales Inventory and Operation Planning (SIOP) using Machine Learning for Root Cause Analysis and Risk Mitigation	444
<i>Sayantani Ghosh, Sai Kumar, Niket Kumar</i>	
Rethinking RobustBench: Is High Synthetic-Test Data Similarity an Implicit Information Advantage Inflating Robustness Scores?.....	448
<i>Chao Pan, Ke Tang, Qing Li, Xin Yao</i>	
Numerically Efficient Parametric Inference for Learning Space-Time Hawkes Processes.....	458
<i>Emilia Siviero, Guillaume Staerman, Stephan Cl��men��on, Thomas Moreau</i>	
Automatic Feature Engineering for Time Series Extrinsic Regression: A Comparative Study of Signal Processing Libraries	468
<i>Aur��lien Renault, Dominique Gay, Nour E. Y. Nair-Benrekia, Vincent Lemaire, Alexis Bondu</i>	
Towards Unified Multimodal Financial Forecasting: Integrating Sentiment Embeddings and Market Indicators via Cross-Modal Attention.....	477
<i>Sarthak Khanna, Armin Berger, David Berghaus, Tobias Deusser, Lorenz Sparrenberg, Rafet Sifa</i>	
Hybrid Federated Learning Framework with Client - Tailored Attentive Feature Extractor for Agricultural Health Monitoring.....	481
<i>Antoni Jaszcz, Agnieszka Polowczyk, Alicja Polowczyk, Katarzyna Wiltos, Dawid Polap, Marcin Wozniak</i>	
Comparative Analysis of Neural Network Architectures for Telemetry-Driven Malware Behavior Classification	488
<i>Michael Schladt, John Gallagher</i>	
Operational Evaluation of Algorithms for Online Streaming Continual Multi-Label Classification	498
<i>Hugo Peuzet, Pascale Kuntz, Frank Meyer, Vincent Lemaire, Killian Le Mau</i>	
Beyond the Pipeline: Analyzing Key Factors in End-To-End Deep Learning for Historical Writer Identification	508
<i>Hanif Rasyidi, Moshiur Farazi</i>	
SpInN: A Time Series Foundation Model for Spectral Indices Recommendation for Earth Observation Applications using Satellite Data	518
<i>Arshveer Kaur, Poonam Goyal, Vansh Bansal, Deep Pandya, Navneet Goyal</i>	
Probabilistic Forecasting Cryptocurrencies Volatility: From Point to Quantile Forecasts.....	528
<i>Grzegorz Dudek, Witold Orzeszko, Piotr Fiszeder</i>	
Optimal Utility Bounds for Differentially Private Gradient Descent in Three-Layer Neural Networks	538
<i>Puyu Wang, Yunwen Lei, Marius Kloft, Yiming Ying</i>	
Metapath-Free Heterogeneous Network Embedding via Heterogeneous Adjacency Matrix Reconstruction Based Mask Graph Autoencoders and Semantic Relation Aggregation	546
<i>Hao-Cheng Ni, Shiou-Chi Li, Jen-Wei Huang</i>	

A Global Dataset-Specific Any-Order Minimal Expectation Baseline for Saliency Scores.....	556
<i>Leonid Schwenke, Martin Atzmueller</i>	
Predicting Player Churn with LLMs: A Comprehensive Evaluation of World Knowledge and Reasoning	566
<i>Tobias Schneider, Lorenz Sparrenberg, Rafet Sifa</i>	
Evaluating Interpretability using Logical Relations: Do Saliency Maps Encode Relevant Information?	576
<i>Leonid Schwenke, Martin Atzmueller</i>	
Diffusion for Explainable Unsupervised Anomaly Detection	588
<i>Elouan Vincent, Alexandre Dréan, Julien Perez, Marc Plantevit, Céline Robardet</i>	
Every Tweet Counts? Sentiment Analysis of Public Response Toward Indonesia's Free Nutritious Meal Program.....	598
<i>Revalda Putawara, Zakiatul Wildani, Nur M. Ulmi, Ayu Mutia, Merysa, Rezky A. Yusuf</i>	
Enhancing Medical Diagnosis with Instance Hardness-Guided Multi-Level Cross-Validation for Imbalanced Learning	607
<i>Mabrouka Salmi, Dalia Atif, Sebastián Ventura</i>	
Multi-Objective Quantum Power System Redispatch	618
<i>Loong K. Lee, Thore Gerlach, Johannes Knaute, Florian Gerhardt, Patrick Völker, Tomislav Maras, Alexander Dotterweich, Nico Piatkowski</i>	
EPDD: Electrocardiogram-Based Pulmonary Disease Detector using Machine Learning.....	630
<i>Hari S. L. V. Vanaparthi, Natasha Interlichia, Xudong Liu, Mona Nasseri, Scott Helgeson</i>	
ModelForge: A Metric Approach to Machine Learning Model Consolidation.....	640
<i>Florian Siepe, Thomas P. Prinz, Thorsten Papenbrock</i>	
Nonnegative Matrix Factorization and the Principle of the Common Cause	652
<i>Edvard Khalafyan, Armen Allahverdyan, Arshak Hovhannisyan</i>	

Author Index