

2024 IEEE 11th International Conference on Data Science and Advanced Analytics (DSAA 2024)

**San Diego, California, USA
6-10 October 2024**



**IEEE Catalog Number: CFP24DSB-POD
ISBN: 979-8-3503-6495-8**

**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:	CFP24DSB-POD
ISBN (Print-On-Demand):	979-8-3503-6495-8
ISBN (Online):	979-8-3503-6494-1
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

Accurate estimation of cross-excitation in multivariate Hawkes process models of infectious diseases	1
<i>Youness Diouane, Frederic Schoenberg and George Mohler</i>	
Masked Contrastive Representation Learning for Self-supervised Visual Pre-training	9
<i>Yuchong Yao, Nandakishor Desai and Marimuthu Palaniswami</i>	
A Novel Particle Swarm Optimization Algorithm for Meta-heuristic Analysis Mechanism Based on Population Learning Strategies and Adaptive Selection of Leadership Particles ..	19
<i>Yuheng Wu, Xingyu Zhu, Wenxu Zhao and Xiaona Xia</i>	
CL-FML: Cluster-based & Label-aware Federated Meta-Learning for On-Demand Classification Tasks	28
<i>Tahani Aladwani, Christos Anagnostopoulos, Shameem Puthiya Parambath and Fani Deligianni</i>	
Diffusion Models for Cross-Domain Image-to-Image Translation with Paired and Partially Paired Datasets	38
<i>Trisk Bell and Dan Li</i>	
Estimate and Reduce Uncertainty in Uncertain Graphs	46
<i>Naheed Anjum Arafat, Ehsan Bonabi Mobaraki, Arijit Khan, Yllka Velaj and Francesco Bonchi</i>	
Efficient Data Completion and Augmentation	58
<i>Antonina Krajewska and Ewa Niewiadomska-Szynkiewicz</i>	
Local Hierarchy-Aware Text-Label Association for Hierarchical Text Classification	68
<i>Ashish Kumar and Durga Toshniwal</i>	
Coded Term Discovery for Online Hate Speech Detection	78
<i>Dhanush Kikkiseti, Raza Ul Mustafa, Wendy Melillo, Roberto Corizzo, Zois Boukouvalas, Jeff Gill and Nathalie Japkowicz</i>	
Data Augmentation's Effect on Machine Learning Models when Learning with Imbalanced Data	88
<i>Damien Dablain and Nitesh Chawla</i>	
Beyond Simple Averaging: Improving NLP Ensemble Performance with Topological-Data-Analysis-Based Weighting	98
<i>Polina Proskura and Alexey Zaytsev</i>	
Deep Contrastive Active Learning for Out-of-domain Filtering in Dialog Systems	106
<i>Roland Oruche, Marcos Zampieri and Prasad Calyam</i>	
Validating Arbitrary Shaped Clusters - A Survey	116
<i>Georg Stefan Schlake and Christian Beecks</i>	
Using Annealing to Accelerate Triangle Inequality k-means	128
<i>Alibek Zhakubayev and Greg Hamerly</i>	

Transparent Neighborhood Approximation for Text Classifier Explanation by Probability-based Editing	137
<i>Yi Cai, Arthur Zimek, Eirini Ntoutsi and Gerhard Wunder</i>	
Multi-view relational evidential c-medoid clustering with adaptive weighted.....	147
<i>Armel Soubeiga, Violaine Antoine and Sylvain Moreno</i>	
Semi-supervised Coarsening of Bipartite Graphs for Text Classification via Graph Neural Network.....	157
<i>Nícolás Roque dos Santos, Diego Minatel, Alan Demétrius Baria Valejo and Alneu de Andrade Lopes</i>	
Let Me Generate That for You: Generative Data Augmentation for Misinformation Detection in Low-Resource Environments	167
<i>Autumn Toney and Lisa Singh</i>	
AI-based Mental Health Assessment for Adolescents Using Their Daily Digital Activities .	178
<i>Do Hyung Kim, Joonsung Lee, Taehwi Lee, Soeun Baek, Seonghyun Jin, Haeun Yoo, Youngeun Cho, Seonghyeon Park, Kwangsu Cho and Chang-Gun Lee</i>	
Improving GNN-based Methods for Scam Detection in Bitcoin Transactions - A Practical Case Study.....	188
<i>Karanjot Saggu, Paula Branco and Guy-Vicent Jourdan</i>	
Model Attribution in LLM-Generated Disinformation: A Domain Generalization Approach with Supervised Contrastive Learning.....	198
<i>Alimohammad Beigi, Zhen Tan, Nivedh Mudiam, Canyu Chen, Kai Shu and Huan Liu</i>	
HeTAN: Heterogeneous Graph Triplet Attention Network for Drug Repurposing.....	208
<i>Farhan Tanvir, Khaled Mohammed Saifuddin, Tanvir Hossain, Arunkumar Bagavathi and Esra Akbas</i>	
Can Large Language Models be Anomaly Detectors for Time Series?	218
<i>Sarah Alnegheimish, Linh Nguyen, Laure Berti-Equille and Kalyan Veeramachaneni</i>	
Discovering Structural Regularities in Time Series via Gaussian Processes	228
<i>Jan David Hüwel and Christian Beecks</i>	
Influence Role Recognition and LLM-based Scholar Recommendation in Academic Social Networks.....	238
<i>Prasad Calyam, Xiyao Cheng, Mayank Kejriwal, Yuanxun Zhang and Lakshmi Srinivas Edara</i>	
Beta k-means: Accelerating k-means Using Probabilistic Cluster Filtering	249
<i>Alibek Zhakubayev and Greg Hamerly</i>	
Media Bias Matters: Understanding the Impact of Politically Biased News on Vaccine Attitudes in Social Media	259
<i>Bohan Jiang, Lu Cheng, Zhen Tan, Ruocheng Guo and Huan Liu</i>	
MetaGAD: Meta Representation Adaptation for Few-Shot Graph Anomaly Detection.....	269
<i>Xiongqiao Xu, Kaize Ding, Canyu Chen and Kai Shu</i>	
More Options for Prelabor Rupture of Membranes, A Bayesian Analysis.....	279
<i>Ashley Klein, Edward Raff, Elisabeth Seamon, Lily Foley and Timothy Bussert</i>	

Human Evaluation of GPT for Scalable Python Programming Exercise Generation	288
<i>Muhammad Fawad Akbar Khan, Max Ramsdell, Ha Nguyen and Hamid Karimi</i>	
Using Hypervectors for Efficient Anomaly Detection in Graph Streams	298
<i>William Arliss, Andrew Godbehere and Graham Mueller</i>	
Forecasting dissolved oxygen based on two-stage decomposition with BiLSTM-Attention and weighted Huber loss function	308
<i>Neha Pant, Durga Toshniwal and Bhola Ram Gurjar</i>	
MicroPPO: Safe Power Flow Management in Decentralized Micro-Grids with Proximal Policy Optimization.....	318
<i>Daniel Ebi, Edouard Fouché, Marco Heyden and Klemens Böhm</i>	
MIS-ME: A Multi-modal Framework for Soil Moisture Estimation	328
<i>Mohammed Rakib, Adil Aman Mohammed, D. Cole Diggins, Sumit Sharma, Jeff Michael Sadler, Tyson Ochsner and Arun Bagavathi</i>	
Magazine supply optimization: a case-study	338
<i>Duong Nguyen, Ana Ulianovici, Sami Achour, Soline Aubry and Nicolas Chesneau</i>	
Predictive Insights into LGBTQ+ Minority Stress: A Transductive Exploration of Social Media Discourse.....	348
<i>Santosh Chapagain, Yuxuan Zhao, Taylor K Rohleen, Shah Muhammad Hamdi, Soukaina Filali Boubrahimi, Ryan E Flinn, Emily M Lund, Dannie Klooster, Jillian R Scheer and Cory J Cascalheira</i>	
PAW: A Deep Learning Model for Predicting Amplitude Windows in Seismic Signals	357
<i>Ariana M. Villegas Suarez, Delaine Reiter, Jonathan Rolfs and Abdullah Mueen</i>	
An Online Calibration Method for Robust Multi-modality 3D Object Detection	367
<i>Xingyu Li, Yige Yao, Jianming Hu and Zhidong Deng</i>	
No Data Left Behind: Exogenous Variables in Long-Term Forecasting of Nursing Staff Capacity	377
<i>Emily Schiller, Simon Müller, Kathrin Ebertsch and Jan-Philipp Steghöfer</i>	
Theory-guided Data-Science for Battery Modeling: Insights from a Comparative Study ...	387
<i>Pawel Bielski, Nico Denner, Simon Bischof, Benedikt Rzepka and Klemens Böhm</i>	
Few-Shot Learning for Detecting Affective States from Keyboard and Mouse Data	397
<i>Ramu Gautam, Beiyu Lin and Mei Yang</i>	
Embedding Ordinality to Binary Loss Function for Improving Solar Flare Forecasting	404
<i>Chetraj Pandey, Anli Ji, Jinsu Hong, Rafal A. Angryk and Berkay Aydin</i>	
Alignment of Multilingual Embeddings to Estimate Job Similarities in Online Labour Market.....	414
<i>Simone D'Amico, Lorenzo Malandri, Fabio Mercorio, Mario Mezzanzanica and Filippo Pallucchini</i>	
Time Series Data Mining: A Unifying View	424
<i>Eamonn Keogh</i>	
Architecture design: from neural networks to foundation models.....	427
<i>Grigorios Chrysos</i>	

Interpretable and Explainable Time Series Mining.....	430
<i>Panagiotis Papapetrou and Zed Lee</i>	
Adversarial Robustness in Graph Neural Networks: Recent Advances and New Frontier...	433
<i>Zhichao Hou, Minhua Lin, Mohamadali Torkamani, Suhang Wang and Xiaorui Liu</i>	
Statistically Inspired Discrepancy Detection for Anomalous Spatio-Temporal Graphs (Extended Abstract)	435
<i>Atharv Tiwari, Shreyash Chatterjee, Siddharth Padmakumar, Sushanth Nair and Bhaskarjyoti Das</i>	
Evaluating the Robustness of Fake News Detectors to Adversarial Attacks with Real User Comments (Extended Abstract).....	437
<i>Annat Koren, Chandler Underwood, Edoardo Serra and Francesca Spezzano</i>	
Novel Discrete Time Series Representation with De Bruijn Graphs for Enhanced Forecasting Using TimesNet.....	439
<i>Mert Cakiroglu, Hasan Kurban, Elham Buxton and Mehmet Dalkilic</i>	
Prompt-Based Contrastive Learning to Combat the COVID-19 Infodemic	442
<i>Zifan Peng, Mingchen Li, Wang Yue and Daniel Y. Mo</i>	
A metaverse platform for air pollution analysis in supporting smart and sustainable city development	443
<i>Juan C. Armijos, Garik Avagyan, Carson K. Leung, Jasmine J. Tabuzo and Aivee F. Teodocio</i>	
Collaborative Intrusion Detection System for Intermittent IoVs Using Federated Learning and Deep Swarm Particle Optimization	453
<i>Farhan Ullah, Gautam Srivastava, Leonardo Mostarda and Diletta Cacciagrande</i>	
A digital twin system to classify tornado climatology data for predicting future disasters..	461
<i>Fengfan Bian, Piers Grenier, Fan Jiang, Carson Leung, Yun Xi Samuel Ning, Adam Pazdor and Hao Ran Harry Pu</i>	
Leveraging Local Data Sampling Strategies to Improve Federated Learning	471
<i>Christoph Düsing, Philipp Cimiano and Benjamin Paaßen</i>	
Human Stone Toolmaking Action Grammar (HSTAG): A Challenging Benchmark for Fine-grained Motor Behavior Recognition	473
<i>Cheng Liu, Xuyang Yan, Zekun Zhang, Cheng Ding, Tianhao Zhao, Shaya Jannati, Cynthia Martinez and Dietrich Stout</i>	
Predicting the survival period of non-small cell lung cancer based on deep learning	481
<i>Guohao Feng, Songyan Han, Qin Wei, Haisheng Hui, Yu Bai, Yongqiang Cheng and Jianxia Liu</i>	
UIE-UnFold: Deep Unfolding Network with Color Priors and Vision Transformer for Underwater Image Enhancement	488
<i>Yingtie Lei, Jia Yu, Yihang Dong, Changwei Gong, Ziyang Zhou and Chi-Man Pun</i>	
PsyPrompt: LLM Prompt Patterns for Goal Contents Pursuit on Social Media	498
<i>Mengli Yu, Yizhun Zhou, Jiahui Wang and Zhenkun Zhou</i>	

Stroke-CVAE-CGAN: a medical imaging data augmentation network incorporating stroke lesion distribution	506
<i>Haisheng Hui, Xueying Zhang, Fenglian Li and Zelin Wu</i>	
An Biometric Model for Iris Images Segmentation and Deep Learning Classification	516
<i>Nawaf Almolhis</i>	
Deep Learning-Based Prediction of Driver Genes in Non-Small Cell Lung Cancer	522
<i>Yu Bai, Songyan Han, Qin Wei, Haisheng Hui, Guohao Feng, Yongqiang Cheng and Jianxia Liu</i>	
Interpretable Deep Learning for Forecasting Online Advertising Costs: Insights from the Competitive Bidding Landscape	530
<i>Fynn Oldenburg, Qiwei Han and Maximilian Kaiser</i>	
SMP-Track: SAM in Multi-Pedestrian Tracking	540
<i>Shiyin Wang, Huafeng Liu, Qiong Wang, Yichao Zhou and Yazhou Yao</i>	
Global-local unified cross-view enhancing framework for person re-identification	549
<i>Tianran Chen, Chen Chen, Yichao Zhou and Xiyuan Hu</i>	
Federated Feature Distillation CLIP for Photovoltaic Panels Defect Detection	558
<i>Yiming Zhong and Min Cao</i>	
Inter-Frame Multiscale Probabilistic Cross-Attention for Surveillance Object Detection ...	565
<i>Huanhuan Xu, Xiyuan Hu and Yichao Zhou</i>	
UPDATE: Mining User-News Engagement Patterns for Dual-Target Cross-Domain Fake News Detection	574
<i>Xuankai Yang, Yan Wang, Xiuzhen Zhang, Shoujin Wang, Huaxiong Wang and Kwok-Yan Lam</i>	
AddShare+: Efficient Selective Additive Secret Sharing Approach for Private Federated Learning	584
<i>Bernard Atiemo Asare, Paula Branco, Iluju Kiringa and Tet Yeap</i>	
Utilizing Structural Metrics from Knowledge Graphs to Enhance the Robustness Quantification of Large Language Models	594
<i>Mohd Ariful Haque, Marufa Kamal, Roy George and Kishor Datta Gupta</i>	
p-ClustVal: A Novel p-adic Approach for Enhanced Clustering and Valuation in High-Dimensional scRNASeq Data	596
<i>Parichit Sharma, Sarthak Mishra, Hasan Kurban and Mehmet Dalkilic</i>	
Towards Personalized Anti-Phishing: Counterfactual Explanation Approach	599
<i>Zhengyang Fan, Wanru Li, Kathryn Laskey and Kuo-Chu Chang</i>	
Rethinking Security: The Resilience of Shallow ML Models	601
<i>Rafael Teixeira, Mário Antunes, João Paulo Barraca, Diogo Gomes and Rui L. Aguiar</i>	