2023 IEEE International Conference on Big Data (BigData 2023)

Sorrento, Italy
15-18 December 2023

Pages 1-689

IEEE Catalog Number: **CFP23BGD-POD**
ISBN: **979-8-3503-2446-4**

Copyright © 2023 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: CFP23BGD-POD
ISBN (Print-On-Demand): 979-8-3503-2446-4
ISBN (Online): 979-8-3503-2445-7

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

The Data Systems Grammar: Self-Designing Systems for the Era of AI 1
Stratos Idreos

Knowledge-Guided Machine Learning: A New Framework for Accelerating Scientific Discovery and Addressing Global Environmental Challenges 2
Vipin Kumar

GPU Acceleration for Data Processing and Analytics 3
Qiong Luo

Have Your Data and Share it Too: Private Data Analytics at Scale 4
Garofalakis Minos

Big-Data Algorithms that Are Not Machine Learning 5
Jeffery Ullman

Solving Tensor Low Cycle Rank Approximation 6
Yichuan Deng, Yeqi Gao, Zhao Song

Subspace Interpolation and Indexing on Stiefel and Grassmann Manifolds as a Lightweight Inference Engine 17
Wenqing Hu, Tiefeng Jiang, Birendra Kathariya, Vikram Abrol, Jiali Zhang, Zhu Li

Mitigating Confounding and Selection Biases in Personalized Recommendation: A Causal Approach 27
Wen Huang, Jingbo Zhou, Xintao Wu, Dejing Dou

CoMABO: Covariance Matrix Adaptation for Bayesian Optimization 37
Hsiang-Yu Ku, Che-Rung Lee

Dynamic Mining Interval to Improve Blockchain Throughput 46
Hou-Wan Long, Xiongfei Zhao, Yain-Whar Si

Online Adaptive Mahalanobis Distance Estimation 56
Lianke Qin, Aravind Reddy, Zhao Song

Efficiently Mining Frequent Representative Motifs in Large Collections of Time Series 66
Stijn J. Rotman, Boris Cule, Len Feremans

Compact Boundary Matrix from a Polytopal Complex for Computing Persistent Homology 76
Rohit P. Singh, Philip A. Wilsey

Time-Varying QoS Estimation Via Non-Negative Latent Factorization of Tensors with Extended Linear Biases 86
Xiuqin Xu, Mingwei Lin, Wenqiang Li, Jianpeng Zhang, Hao Wu

Big Data Management System Architectures: From Opportunities to Challenges [Vision] 96
Radhika Joshi, Calvin Thomas Dani, Shiva Jahangiri

Ocean Data Quality Assessment Through Outlier Detection-Enhanced Active Learning 102
Na Li, Yiyang Qi, Ruyue Xin, Zhiming Zhao

Verifying Relational Explanations: A Probabilistic Approach 108
Abisha Thapa Magar, Anup Shakya, Somdeb Sarkhely, Deepak Venugopal

Vision Paper: Hyperspectral Analysis of Finger Skin Reflectance for Resilient Biometric Systems 116
Emanuela Marasco

KF K-Means: A High Performance K-Means Implementation Using Kernel Fusion 121
Kaiming Ouyang, Vincent Tran, Jinyang Liu, Bryan M. Wong, Zizhong Chen

Fast Heavy Inner Product Identification Between Weights and Inputs in Neural Network Training 128
Lianke Qin, Saayan Mitra, Zhao Song, Yuanyuan Yang, Tianyi Zhou

Quality Ranking for Synthetically Generated Images 134
Ali Sekmen, Bahadir Bilgin, Guy Sereff

HARVEST: High-Performance Artificial Vision Framework for Expert Labeling Using Semi-Supervised Training 139
Nawras Alnaasan, Matthew Lieber, Aamir Shafi, Hari Subramoni, Scott Shearer, Dhabaleswar K Panda

Computing Group-By and Aggregates on Massively Parallel Systems 149
Chengcheng Mou, Minh Pham, Benjamin Hsu, Yicheng Tu

User-Guided Page Merging for Memory Deduplication in Serverless Systems 159
Wei Qiu, Marcin Copik, Yun Wang, Alexandru Calotoiu, Torsten Hoefler

PieRank: Embedded Large-Scale Sparse Matrix Processing 170
Michelle S. Zhou

Predicting Dynamic Memory Requirements for Scientific Workflow Tasks 182
Jonathan Bader, Nils Diedrich, Lauritz Thamsen, Odej Kao

Foundation Data Space Models: Bridging the Artificial Intelligence and Data Ecosystems (Vision Paper) 190
Edward Curry, Tarek Zaarour, Yang Yang, Mohan Timilsina, Majjed Al-Qatf, Rafiqul Haque

Towards Open World Event Processing: A Paradigm and Research Agenda for Event Processing for the Internet of Multimedia Things (Vision Paper) 196
Edward Curry

ChimeraPy: A Scientific Distributed Streaming Framework for Real-Time Multimodal Data Retrieval and Processing 201
Eduardo Davalos, Umesh Timalsina, Yike Zhang, Jiayi Wu, Joyce Horn Fonteles, Gautam Biswas

Co-Tuning of Cloud Infrastructure and Distributed Data Processing Platforms 207
Isuru Dharmadasa, Faheem Ullah

On Overcoming HPC Challenges of Trillion-Scale Real-World Graph Datasets 215
Mohsen Koohi Esfahani, Paolo Boldi, Hans Vandierendonck, Peter Kilpatrick, Sebastiano Vigna

Distributed Quantum Learning with co-Management in a Multi-Tenant Quantum System 221
Anthony D'Onofrio, Amir Hossain, Lesther Santana, Naseem Machlovi, Samuel Stein, Jinwei Liu, Ang Li, Ying Mao

Scientific Error-Bounded Lossy Compression with Super-Resolution Neural Networks 229
Jinyang Liu, Sheng Di, Sian Jin, Kai Zhao, Xin Liang, Zizhong Chen, Franck Cappello

Sailfish: A Dependency-Aware and Resource Efficient Scheduling for Low Latency in Clouds..... 237
Jinwei Liu, Yingjie Lao, Ying Mao, Rajkumar Buyya

Performance of Null Handling in Array Databases 247
Dimitar Mišev, Mikhail Rodionychev, Peter Baumann

Accelerating Multilingual Applications with In-Memory Array Sharing..... 255
Masanobu Nozawa, Satoshi Imamura, Kenji Kono

GSPICE: Model-Based Event Shedding in Complex Event Processing 263
Ahmad Slo, Sukanya Bhowmik, Kurt Rothermel

Forgetting User Preference in Recommendation Systems with Label-Flipping..... 271
Manal A. Alshehri, Xiangliang Zhang

On-Street Parking Prediction: A Comparative Study 282
Stefano Fiorini, Michele Ciavotta, Carlo Liberto, Gaetano Valenti

Detecting Semantic Errors in Tables Using Textual Evidence 292
Minh Pham, Craig A. Knoblock, Muhao Chen

TRANSGNN: Towards Knowledge Enhanced Top-K Recommendation Via Transformed Heterogeneous Graph Neural Network 304
Shen Wang, Ziwei Fan, Jibing Gong, Xiaokai Wei, Philip S. Yu

Large-Scale User Preference Tracking Via Asynchronous and Asymmetric Updating at Twitter..... 315
Ga Wu, Shivam Khare, Li Zhou, Yael Brumer, Jun-Ping Ng, Ruowei Wang

Multiloss-Based Optimization for Time Series Data Augmentation 325
Omar Bahri, Peiyu Li, Soukaïna Filali Boubrahimi, Shah Muhammad Hamdi

Efficient Missing Value Imputation by Maximum Distance Likelihood..... 331
Savong Bou, Toshiyuki Amagasa, Hiroyuki Kitagawa, Salman Ahmed Shaikh, Akiyoshi Matono

Self-Configured Entity Resolution with pyJedAI..... 339
Vasilis Efthymiou, Ekaterini Ioannou, Manos Karvounis, Manolis Koubarakis, Jakub Maciejewski, Konstantinos Nikoletos, George Papadakis, Dimitris Skoutas, Yannis Velegrakis, Alexandros Zeakis

A Data Integration Framework with Multi-Source Big Data for Enhanced Forest Fire Prediction..... 344
Parveen Kaur, Kshirasagar Naik, Richard Purcell, Srinivas Sampalli, Chung-Horng Lung, Marzia Zaman, Abdul Mutakabbir

Leveraging Uncertainty Quantification for Reducing Data for Recommender Systems..... 352
Xi Niu, Ruhani Rahman, Xiangcheng Wu, Zhe Fu, Depeng Xu, Riyi Qiu

Multi-Layer Caching and Parallel Streaming for Large Scale Cloud Optimized Point Cloud Data Visualization Using WebGPU 360
Pravin Poudel, Will Usher, Steve Petruzza

An Analysis of Combined Data Augmentation in Time Series Prediction Tasks Using Discrete Wavelet Multilevel Decomposition 366
Yuewen Shen, Xinhao Yang, Jiayu Zhao, Ze Li

Fine-Grained Urban Population Distribution Estimation Using Image Super-Resolution Model with Rich Auxiliary Information 372
Naoto Takeda, Akihiro Kobayashi, Yudai Yamazaki, Daisuke Kamisaka

Dual-Path Side Information Fusion for Sequential Recommendation 377
Yu Zhang, Haiwei Pan, Kejia Zhang, Tianming Zhang, Qingquan Ren, Wenjie Li

A Scalable Approach to Aligning Natural Language and Knowledge Graph Representations: Batched Information Guided Optimal Transport 383
Alexander Kalinowski, Deepayan Datta, Yuan An

Group-Aware Interest Disentangled Dual-Training for Personalized Recommendation 393
Xiaolong Liu, Liangwei Yang, Zhiwei Liu, Xiaohan Li, Mingdai Yang, Chen Wang, Philip S. Yu

Cross-Domain Fake News Detection Based on Coarse-Fine Grained Environments Reflecting Public Expectation 403
Guojun Liu, Yuefeng Ma, Xun Liang

Pandemic Culture Wars: Partisan Differences in the Moral Language of COVID-19 Discussions 413
Ashwin Rao, Siyi Guo, Sze Yuh Nina Wang, Fred Morstatter, Kristina Lerman

Discovery and Matching Numerical Attributes in Data Lakes 423
Pattara Sukprasert, Gromit Yeuk-Yin Chan, Ryan A. Rossi, Fan Du, Eunyee Koh

MERMAID: A Dataset and Framework for Multimodal Meme Semantic Understanding 433
Shaun Toh, Adriel Kuek, Wen-Haw Chong, Roy Ka-Wei Lee

Sketching Multidimensional Time Series for Fast Discord Mining 443
Chin-Chia Michael Yeh, Yan Zheng, Menghai Pan, Huiyuan Chen, Zhongfang Zhuang, Junpeng Wang, Liang Wang, Wei Zhang, Jeff M. Phillips, Eamonn Keogh

Multi-State Brain Network Discovery 453
Hang Yin, Yao Su, Xinyue Liu, Thomas Hartvigsen, Yanhua Li, Xiangnan Kong

Social Event Detection with Reinforced Deep Heterogeneous Graph Attention Network 463
Yongsheng Yu, Jia Wu, Jian Yang

ID-MixGCL: Identity Mixup for Graph Contrastive Learning 473
Gehang Zhang, Bowen Yu, Jiangxia Cao, Xinghua Zhang, Jiawei Sheng, Chuan Zhou, Tingwen Liu

Dual-Teacher Knowledge Distillation for Strict Cold-Start Recommendation 483
Weizhi Zhang, Liangwei Yang, Yuwei Cao, Ke Xu, Yuanjie Zhu, Philip S. Yu

Towards Bi-Level Out-Of-Distribution Logical Reasoning on Knowledge Graphs 493
Fatimah Alotaibi, Dawei Zhou

SuperBlocking: An Efficient Blocking Technique for Record Linkage 498
Joyanta Basak, Sartaj Sahni, Sanguthevar Rajasekaran

Chain-Based Outlier Detection for Complex Data Scenarios 504
Huiwen Dong, Qing-Guo Wang, Wei Ding

Mining Large-Scale Knowledge Graphs for Chemical Reaction Fingerprints 510
Blake B. Gaines, Minghu Song, Jinbo Bi

Constructing Knowledge Graph from Cyber Threat Intelligence Using Large Language Model 516
Jiehui Liu, Jieyu Zhan

Maximum Flow on Highly Dynamic Graphs 522
Juntong Luo, Scott Sallinen, Matei Ripeanu

Increase Merge Efficiency in LSM Trees Through Coordinated Partitioning of Sorted Runs 530
Qizhong Mao, Vagelis Hristidis

PSQS: Parallel Semantic Querying Service for Self-Describing File Formats..... 536
Chenxu Niu, Wei Zhang, Suren Byna, Yong Chen

Efficient Estimation of Local Causal Effects in Graphs Via Neighborhood Pooling 542
Ryan Rossi, Somdeb Sarkhel, Nesreen Ahmed

Integrating Noisy Knowledge into Language Representations for E-Commerce Applications..... 548
Karan Samel, Jun Ma, Zhengyang Wang, Tong Zhao, Irfan Essa

CGP: Centroid-Guided Graph Poisoning for Link Inference Attacks in Graph Neural Networks 554
Haozhe Tian, Haibo Hu, Qingqing Ye

K-Bitruss-Attributed Weighted Community Search on Attributed Weighted Bipartite Graphs..... 562
Zuxuan Zhang, Hongxi Li, Kai Huang, Yuncheng Jiang

An Adaptive Hierarchical Method for Anytime Set-Wise Clustering of Variable and High-Speed Data Streams..... 568
Aarti, Jagat Sesh Challa, Utkarsh Darolia, Mridul Chandak, Poonam Goyal, Navneet Goyal

NeuEmot: Mitigating Neutral Label and Reclassifying False Neutrals in the 2022 FIFA World Cup Via Low-Level Emotion 578
Ademola Adesokan, Sanjay Madria

mTADS: Multivariate Time Series Anomaly Detection Benchmark Suites 588
David Baumgartner, Helge Langseth, Heri Ramampiaro, Kenth Engø-Monsen

Improving Knowledge Distillation for Federated Learning on Non-IID Data 598
Zongyi Chen, Sanchuan Guo, Liyan Shen, Xi Zhang, Zhuonan Chang

Difforecast: Image Generation Based Highway Traffic Forecasting with Diffusion Model 608
Pengnan Chi, Xiaoliang Ma

WellFactor: Patient Profiling Using Integrative Embedding of Healthcare Data 616
Dongjin Choi, Andy Xiang, Ozgur Ozturk, Deep Shrestha, Barry Drake, Hamid Haidarian, Faizan Javed, Haesun Park

Density-Based Prototypical Contrastive Learning on Visual Representations 626
Yueh-Lin Chung, Hsing-Kuo Pao

Aspect-Based Meeting Transcript Summarization: A Two-Stage Approach with Weak Supervision on Sentence Classification..... 636
Zhongfen Deng, Seunghyun Yoon, Trung Bui, Franck Dernoncourt, Quan Hung Tran, Shuaiqi Liu, Wenting Zhao, Tao Zhang, Yibo Wang, Philip S. Yu

Feature Selection on Big Data Using Masked Sparse Bottleneck Centroid-Encoder..... 646
Tomojit Ghosh, Michael Kirby

Interactive Visualization for Smart Power Grid Efficiency and Outage Exploration 656
Dongyun Han, Isaac Cho

Striking a Balance in Fairness for Dynamic Systems Through Reinforcement Learning 662
Yaowei Hu, Jacob Lear, Lu Zhang

Adaptive Shrinkage Denoising and Sequential State Extraction Model for Vibration Event Recognition 672
Wanchang Jiang, Yuxin Jiang

Integrating Staleness and Shapley Value Consistency for Efficient K-Asynchronous Federated Learning 680
Yuhui Jiang, Xingjian Lu, Wei Mao, Ying Lin

Towards Cost-Effective and Resource-Aware Aggregation at Edge for Federated Learning 690
Ahmad Faraz Khan, Yuze Li, Xinran Wang, Sabaat Haroon, Haider Ali, Yue Cheng, Ali R. Butt, Ali Anwar

Spatial-Temporal Dynamic Graph Diffusion Convolutional Network for Traffic Flow Forecasting 700
Yang Li, Yaying Zhang

RSP-GcForest: A Distributed Deep Forest Via Random Sample Partition 710
Mark Junjie Li, Wenzhu Cai, Yigang Lin, Sunjie Huang, Joshua Zhexue Huang, Patrick Xiaogang Peng

CELS: Counterfactual Explanations for Time Series Data Via Learned Saliency Maps 718
Peiyu Li, Omar Bahri, Soukaïna Filali Boubrahimi, Shah Muhammad Hamdi

SEED: An Effective Model for Highly-Skewed Streamflow Time Series Data Forecasting 728
Yanhong Li, Jack Xu, David C. Anastasiu

Unveiling the Road Matrix: Development of Versatile Pavement Surface Models from Low-Resolution Aerial Images 738
Ling-Feng Liao, Deeksha Arya, Yoshihide Sekimoto

Long-Term Time Series Forecasting Based on Decomposition and Neural Ordinary Differential Equations 748
Seonkyu Lim, Jaehyeon Park, Seojin Kim, Hyowon Wi, Haksoo Lim, Jinsung Jeon, Jeongwhan Choi, Noseong Park

Semi-Supervised 3D Video Information Retrieval with Deep Neural Network and Bi-Directional Dynamic-Time Warping Algorithm 758
Yintai Ma, Diego Klabjan

Scalable Homology Classification Through Decomposed Euler Characteristic Curves 768
Nicholas O. Malott, Philip A. Wilsey

Heterogeneous Domain Adaptation with Positive and Unlabeled Data 778
Junki Mori, Ryo Furukawa, Isamu Teranishi, Jun Sakuma

Generalized Causal Tree for Uplift Modeling 788
Preetam Nandy, Xiufan Yu, Wanjun Liu, Ye Tu, Kinjal Basu, Shaunak Chatterjee

Multi-View Deep Markov Models for Time Series Anomaly Detection 799
Phuong Nguyen, Hiep Tran, Tuan M. V. Le

A Novel Training Mechanism for Health Indicator Construction and Remaining Useful Lifetime (RUL) Prediction 809
Hanbyeol Park, Dohee Kim, Minseop Kim, Mingyu Park, Hyerim Bae, Yunkyung Park

Hybrid Loss for Hierarchical Multi–label Classification Network 819
Wenting Qi, Charalampos Chelmis

OrthoNets: Orthogonal Channel Attention Networks 829
Hadi Salman, Caleb Parks, Matthew Swan, John Gauch

SST: Semantic and Structural Transformers for Hierarchy-Aware Language Models in E-Commerce 838
Karan Samel, Houyu Zhang, Jun Ma, Haoming Jiang, Qing Ping, Sheng Wang, Yi Xu, Belinda Zeng, Trishul Chilimbi

Interpretable Sock Puppet Attribution 847
Chun Wei Seah, Chia-Yu Hung, Yujia Hu, Hai Leong Chieu, Roy Ka-Wei Lee

Multi-Instance Bias Suppression for Enhanced Generalization in Breast Cancer Diagnosis : Harnessing Histopathological Big Data Insights 857
Syed Attique Shah, Xiaoyang Zeng, Awais Ahmed, Shaheed Parvez, Rui Xi, Mengshu Hou

ADAMM: Anomaly Detection of Attributed Multi-Graphs with Metadata: A Unified Neural Network Approach 865
Konstantinos Sotiropoulos, Lingxiao Zhao, Pierre Jinghong Liang, Leman Akoglu

Leveraging Large Language Models for Structure Learning in Prompted Weak Supervision 875
Jinyan Su, Peilin Yu, Jieyu Zhang, Stephen H. Bach

ABF-FNN: A New Fuzzy Neural Network for Predicting Coal Mine Gas Concentration Hazard 885
Yimin Sun, Xiaobo Zhang, Zhehao Zhang, Yunyang Wu, Haihao Tang, Haonan Luo

RoadFormer: Road-Anchored Adversarial Dynamic Graph Transformer for Unlimited-Range Traffic Incident Impact Prediction 895
Yanshen Sun, Kaiqun Fu, Chang-Tien Lu

Scalable Volume Visualization for Big Scientific Data Modeled by Functional Approximation 905
Jianxin Sun, David Lenz, Hongfeng Yu, Tom Peterka

Unvisited Out-Of-Town POI Recommendation with Simultaneous Learning of Multiple Regions 915
Rikuto Tsubouchi, Takahiro Hara, Kei Yonekawa, Shuichiro Haruta

Flexible Communication for Optimal Distributed Learning Over Unpredictable Networks 925
Sahil Tyagi, Martin Swany

IIP-Transformer: Intra-Inter-Part Transformer for Skeleton-Based Action Recognition 936
Qingtian Wang, Shuze Shi, Jiabin He, Jianlin Peng, Tingxi Liu, Renliang Weng

A Prior-Guided Meta-Weight Network for Long-Tailed Noisy Data in Item Categorization 946
Tianqi Wang, Lei Chen, Jing Gao

GNN-Ensemble: Towards Random Decision Graph Neural Networks 956
Wenqi Wei, Mu Qiao, Divyesh Jadav

Roughness Index for Loss Landscapes of Neural Network Models of Partial Differential Equations* 966
Keke Wu, Xiangru Jian, Rui Du, Jingrun Chen, Xiang Zhou

Ultra-Wide Band Positioning with Signal Interference Based on Two-Stream Residual Network 976
Xueming Xu, Ruirui Zhao, Jichao Li, Jianbin Sun

Tripletformer for Probabilistic Interpolation of Irregularly Sampled Time Series 986
Vijaya Krishna Yalavarthi, Johannes Burchert, Lars Schmidt-Thieme

Traffic Demand Prediction Based on Multi-Dimensional Graph Convolutional Network..... 996
Peiying Zeng, Liying Jiang, Yongxuan Lai, Fan Yang

Document-Level Event Argument Extraction Based on Bidirectional Span Detection..... 1005
Yong Zhang, Feng Xiong, Kai Zhang, Jiaxian Wang

Enhancing Representation Learning with Label Association for Multi-Label Text Classification..... 1012
Wang Zhang, Xin Wang, Yuhong Wu, Xingpeng Zhang, Huayi Zhan

Label Freedom: Stable Diffusion for Remote Sensing Image Semantic Segmentation Data Generation 1022
Chenbo Zhao, Yoshiki Ogawa, Shenglong Chen, Zhehui Yang, Yoshihide Sekimoto

CASTLE: A Cascaded Spatio-Temporal Approach for Long-Lead Streamflow Forecasting..... 1031
Yong Zhuang, David L. Small, Patrick D Flynn, Wahid Palash, Shafiqul Islam, Ping Chen, Wei Ding

OLR-WA: Online Weighted Average Linear Regression in Multivariate Data Streams..... 1039
Mohammad Abu-Shaira, Alejandro Rodriguez, Greg Speegle, Victor Sheng, Ishfaq Ahmad

Self-Supervision for Tackling Unsupervised Anomaly Detection: Pitfalls and Opportunities 1047
Leman Akoglu, Jaemin Yoo

M2-Mixer: A Multimodal Mixer with Multi-Head Loss for Classification from Multimodal Data 1052
Grigor Bezirganyan, Sana Sellami, Laure Berti-équille, Sébastien Fournier

Learning with Variational Objectives on Measures 1059
Vivien Cabannes

Improve Interpretability of Information Bottlenecks for Attribution with Layer-Wise Relevance Propagation..... 1064
Xiongren Chen, Jiuyong Li, Jixue Liu, Stefan Peters, Lin Liu, Thuc Duy Le, Anthony Walsh

Toward Efficient and Incremental Spectral Clustering Via Parametric Spectral Clustering..... 1070
Jo-Chun Chen, Hung-Hsuan Chen

Hypergraph Neural Networks for Time-Series Forecasting 1076
Hongjie Chen, Ryan A. Rossi, Kanak Mahadik, Sungchul Kim, Hoda Eldardiry

Federated Learning for Sparse Principal Component Analysis 1081
Sin Cheng Ciou, Pin Jui Chen, Elvin Y. Tseng, Yuh-Jye Lee

JPAVE: A Generation and Classification-Based Model for Joint Product Attribute Prediction and Value Extraction 1087
Zhongfen Deng, Hao Peng, Tao Zhang, Shuaiqi Liu, Wenting Zhao, Yibo Wang, Philip S. Yu

Drift-Aware Anomaly Detection for Non-Stationary Time Series..... 1095
Yang Gao, Ying Li, Yang Li, Zunlei Feng, Mingli Song, Xingyu Wang, Chun Chen

Autoencoders and Generative Adversarial Networks for Imbalanced Sequence Classification.....1101
Stephanie Ger, Yegna Subramanian Jambunath, Diego Klabjan

Cohesive Attention-Based Explanations for Sequences and Explainability in Presence of Event Types 1109
Stephanie Ger, Yegna Subramanian Jambunath, Diego Klabjan, Jean Utke

LogGPT: Log Anomaly Detection Via GPT 1117
Xiao Han, Shuhan Yuan, Mohamed Trabelsi

Transferring Lottery Tickets in Computer Vision Models: A Dynamic Pruning Approach 1123
Qisheng He, Ming Dong

Learning Pruned Structure and Weights Simultaneously from Scratch: An Attention Based Approach 1129
Qisheng He, Weisong Shi, Ming Dong

Causal Associations Between Temporal Events 1135
Nguyen Ho, Trinh Cong Le, Van Long Ho, Nguyen Tuong Huynh, Linh Ngo

Quantitatively Evaluating the Validity of Contrastive Generators for Recourse 1143
Ian Howell, Eleanor Quint, Hongfeng Yu

Contrastive Multi-View Learning for Graph Stucture Learning 1151
Kai Huang, Dengzhe Liang, He Wang, Yuncheng Jiang

Deep Learning for Spatiotemporal Big Data: Opportunities and Challenges [Vision Paper] 1157
Zhe Jiang

Temporal Knowledge Graph Question Answering Models Enhanced with GAT 1162
Wenjuan Jiang, Yi Guo, Jiaojiao Fu

RAxC: Reflexivity-Based Approximate Computing Techniques for Efficient Remote Sensing 1168
Aaditi Kapre, Shruti Kunde, Sparsh Mittal, Rekha Singhal

Label Hierarchy Alignment for Improved Hierarchical Text Classification 1174
Ashish Kumar, Durga Toshniwal

CTT: Causally Informed Tensor Train Decomposition 1180
Mao-Lin Li, K Selçuk Candan, Maria Luisa Sapino

Feature Acquisition Using Monte Carlo Tree Search* 1188
Sungsoo Lim, Diego Klabjan, Mark B. Shapiro

VHGN: Cooperative Multiagent Reinforcement Learning with Historical Information Graph 1194
Jiahao Liu, Jiale Xu, Haijia Bao, Ya Li

A Hybrid Model Based on Multi-Level External Knowledge for Chinese Semantic Matching 1200
Haoyang Ma, Hongyu Guo

A Robust Classifier Under Missing-Not-At-Random Sample Selection Bias 1206
Huy Mai, Wen Huang, Wei Du, Xintao Wu

CSER: Enhancing Cybersecurity Entity Recognition Through Multidimensional Feature Fusion 1214
Md Abu Marjan, Toshiyuki Amagasa

DISCERN: Leveraging Knowledge Distillation to Generate High Resolution Soil Moisture Estimation from Coarse Satellite Data 1222
Abdul Matin, Paahuni Khandelwal, Shrideep Pallickara, Sangmi Lee Pallickara

Optimal Sampling for Moving Object Trajectory Tracking in Smart Transportation Systems: A Transformer-Based Approach 1230
Usman Shuaibu Musa, Yifan Guo, Cheng Qiant, Wei Yu

Prompting Large Language Models for Topic Modeling 1236
Han Wang, Nirmalendu Prakash, Nguyen Khoi Hoang, Ming Shan Hee, Usman Naseem, Roy Ka-Wei Lee

Ego-Network Transformer for Subsequence Classification in Time Series Data 1242
Chin-Chia Michael Yeh, Huiyuan Chen, Yujie Fan, Xin Dai, Yan Zheng, Vivian Lai, Junpeng Wang, Zhongfang Zhuang, Liang Wang, Wei Zhang, Eamonn Keogh

Detecting and Classifying Changes in Traffic Rules Using Induction Loop Data 1248
Aziza Zhanabatyrova, Clayton Leite, Yu Xiao

Humor Level Recognition Based on Prompt Learning and Contrastive Learning 1256
Shaowu Zhang, Xijun Cui, Tongyue Zhang, Hongfei Lin, Liang Yang

Spatial-Temporal Flow Holistic Interaction Graph Convolution Network for Bidirectional Traffic Flow Forecasting 1262
Canyang Zhang, Qi Zheng, Yaying Zhang

FOCAL: A Cost-Aware Video Dataset for Active Learning 1269
Kiran Kokilepersaud, Yash-Yee Logan, Ryan Benkert, Chen Zhou, Mohit Prabhushankar, Ghassan Alregib, Enrique Corona, Kunjan Singh, Mostafa Parchami

Robust Fraud Detection Via Supervised Contrastive Learning 1279
M. S Vinay, Shuhan Yuan, Xintao Wu

Measuring Bias 1289
Aida Sharif Rohani, Ricardo Baeza-Yates

Improving Adversarially Robust Sequential Recommendation Through Generalizable Perturbations 1299
Xun Yao, Ruyi He, Xinrong Hu, Jie Yang, Yi Guo, Zijian Huang

Vision Paper: Leveraging Industrial Big Data – Past, Present, and Future of the World Wide Lab 1308
Mohamed Behery, Felix Glawe, István Koren, Martina Ziefle, Gerhard Lakemeyer, Philipp Brauner

Predicting Linguistically Sophisticated Social Determinants of Health Disparities with Neural Networks: The Case of LGBTQ+ Minority Stress 1314
Cory J. Cascalheira, Santosh Chapagain, Ryan E. Flinn, Yuxuan Zhao, Soukaina Filali Boubrahimi, Dannie Klooster, Alejandra Gonzalez, Emily M. Lund, Danica Laprade, Jillian R. Scheer, Shah Muhammad Hamdi

Qobra: Fast Verification of Transactional Serializability with Quantum Annealing 1322
Natsuki Hamada, Kazuhiro Saito, Hideyuki Kawashima

Zooming into Abnormal Events in Video Conferencing 1330
Shmuel Horowitz, Dima Kagan, Galit Fuhrmann Alpert, Michael Fire

A Privacy-Preserving Semi-Decentralized Personalized Recommendation System 1336
Carson K. Leung, Evan W. R. Madill, Qi Wen

Detecting Compromised IoT Devices Using Autoencoders with Sequential Hypothesis Testing 1344
Md Mainuddin, Zhenhai Duan, Yingfei Dong

Improving the Utility of Differentially Private SGD by Employing Wavelet Transforms 1352
Kanishka Ranaweera, David Smith, Dinh C. Nguyen, Pubudu N. Pathirana, Ming Ding, Thierry Rakotoarivelo, Aruna Seneviratne

An Exploratory Study of Vulnerabilities in Big Data Systems..... 1360
Nikolas Tyllis, Faheem Ullah, Muhammad Uzair

Analyzing Threat Vectors in ICS Cyberattacks 1368
Ismael Villanueva-Miranda, Monika Akbar

Active Data Reconstruction Attacks in Vertical Federated Learning..... 1374
Minh N. Vu, Tre' R. Jeter, Raed Alharbi, My T. Thai

PGAN-KD:Member Privacy Protection of GANs Based on Knowledge Distillation..... 1380
Tianhan Zhang, Juan Yu, Jianmin Han, Hao Peng, Sheng Qiu

Causal Fairness-Guided Dataset Reweighting Using Neural Networks..... 1386
Xuan Zhao, Klaus Broelemann, Salvatore Ruggieri, Gjergji Kasneci

M^2–CNN: A Macro-Micro Model for Taxi Demand Prediction 1395
Shih-Fen Cheng, Prabod Rathnayaka

Combining Demographic Tabular Data with BERT Outputs for Multilabel Text Classification in Higher Education Survey Data 1403
Kevin Chovanec, John Fields, Praveen Madiraju

DRLO: Deep Representation Learning for Large Scale Off-Track Satellite Remote Sensing Data..... 1410
Xin Huang, Chenxi Wang, Wenbin Zhang, Sanjay Purushotham, Jianwu Wang

Predicting Impression Evaluation of Building Exterior Appearance Using Street Image Big Data and Deep Learning 1419
Yusuke Imadegawa, Takuya Oki, Yoshiki Ogawa, Chenbo Zhao

Circular Arc Length-Based Kernel Matrix for Protein Sequence Classification 1429
Taslim Murad, Sarwan Ali, Prakash Chourasia, Haris Mansoor, Murray Patterson

FNDaaS: Content-Agnostic Detection of Websites Distributing Fake News 1438
Panagiotis Papadopoulos, Dimitris Spythouris, Evangelos P. Markatos, Nicolas Kourtellis

Multimodal Co-Attention Transformer for Video-Based Personality Understanding 1450
Mingwei Sun, Kunpeng Zhang

Finformer: A Static-Dynamic Spatiotemporal Framework for Stock Trend Prediction..... 1460
Yi Zu, Jiacong Mi, Lingning Song, Shan Lu, Jieyue He

Development of a Computer Vision-Based Road Physical Feature Extraction 1470
Juliana Marie Agulto, Yeohan Lorenzo Noroña, Gian Carlo Gutierrez, Eric Benabaye, Geanne Franco, Ann Franchesca Laguna, Jocelynn Cu, Joel Ilao, Macario Cordel

Zero-Shot Learning with Minimum Instruction to Extract Social Determinants and Family History from Clinical Notes Using GPT Model 1476
Neel Jitesh Bhate, Ansh Mittal, Zhe He, Xiao Luo

A Novel Approach for Deep Learning-Powered Forecasting of Market Bottoms in Cryptocurrency and Stock Trading..... 1481
D. M. D. K. Dasanayake, H. Y. Dilshan, H. D. K. Y. Rathnaweera, Sapumal Ahangama, Indika Perera

MobiSpaces: An Architecture for Energy-Efficient Data Spaces for Mobility Data 1487
Christos Doulkeridis, Georgios M. Santipantakis, Nikolaos Koutroumanis, George Makridis, Vasilis Koukos, George S. Theodoropoulos, Yannis Theodoridis, Dimosthenis Kyriazis, Pavlos Kranas, Diego Burgos, Ricardo Jimenez-Peris, Mariana M G Duarte, Mahmoud Sakr, Esteban Zimányi, Anita Graser, Clemens Heistracher, Kristian Torp, Ioannis Chrysakis, Theofanis Orphanoudakis, Evgenia Kapassa, Marios Touloupou, Jürgen Neises, Petros Petrou, Sophia Karagiorgou, Rosario Catelli, Domenico Messina, Marcelo Corrales Compagnucci, Matteo Falsetta

A Hybrid Method of Sentiment Analysis and Machine Learning Algorithm for the U.S. Presidential Election Forecasting 1495
Guocheng Feng, Kaihao Chen, Huaiyu Cai, Zhijian Li

Comparing ML OCR Engines on Texts from 19th Century Written in the Romanian Transitional Script 1501
Marc Frincu, Marius E. Penteliuc, Simina Frincu, Gheorghe Bran, Manuela Zanescu

Exploring Dementia Risk in Cardiovascular Patients Through AI-Enabled Big Data Analysis 1507
Shreya Iyer, Jacob Brain, Reza Omidvar, Mario Siervo, Nadeem Qureshi, Ralph Kwame Akyea, Stephen Kaptoge, Grazziela Figueredo, Blossom Stephan

Low-Latency Visual Previews of Large Synchrotron Micro-CT Datasets 1513
Nicholas Tan Jerome, Suren Chilingaryan, Thomas Van De Kamp, Andreas Kopmann

Interpretable Solar Flare Prediction with Sliding Window Multivariate Time Series Forests 1519
Anli Ji, Berkay Aydin

An Interactive Framework of Balancing Evaluation Cost and Prediction Accuracy for Knowledge Tracing 1525
Weili Sun, Boyu Gao, Jiahao Chen, Shuyan Huang, Weiqi Luo

A Big Data Approach to Augmenting the Huff Model with Road Network and Mobility Data for Store Footfall Prediction 1533
Ming Hui Tan, Kar Way Tan, Hoong Chuin Lau

Tabular Data to Image Generation: Benchmark Data, Approaches, and Evaluation 1539
Alex Tang, Gromit Chan, Ryan A. Rossi, Chang Xiao, Eunyee Koh

SASTDGCN: Self-Attention Based Spatial-Temporal Double Graph Convolutional Networks for Traffic Flow Forecasting 1544
Jingjia Wan, Yu Wu

YOLOv5s-BSS: A Novel Deep Neural Network for Crack Detection of Road Damage 1550
Conghua Wei, Qianjun Zhang, Xiaobo Zhang, Yan Yang, Jixin Zhang, Donghai Zhai

AQUA: A Framework for Spatiotemporal Analysis and Visualizations of Water Quality Data at Scale 1555
Matthew Young, Sangmi Pallickara, Shrideep Pallickara

ATADA: Adaptive Time Aware Anomaly Detection Approach for Real-Time Intelligent Transportation Systems 1563
Ke Zhang, Hengchang Liu, Siobhán Clarke

Audience Prospecting for Dynamic-Product-Ads in Native Advertising 1571
E. Abutbul, Y. Kaplan, N. Krasne, O. Somekh, O. David, O. Duvdevany, E. Segal

Refactoring ETL Flows in the Wild 1581
Dolev Adas, Ohad Eytan, Guy Khazma, Josep Sampé, Paula Ta-Shma

HungerGist: An Interpretable Predictive Model for Food Insecurity 1591
Yongsu Ahn, Muheng Yan, Yu-Ru Lin, Zian Wang

VEMUT-Sub-Sampling: A Novel Method for Sparse Multivariate Time-Series Vehicle Data 1601
Nadin-Katrin Apel, Constantinos Antoniou

Efficient Extraction of Insights at the Edges of Distributed Systems 1610
Amadou Ba, Fearghal O'Donncha, Joern Ploennigs, Muneeza Azmat

AI ATAC 1: An Evaluation of Prominent Commercial Malware Detectors 1620
Robert A. Bridges, Brian Weber, Justin M. Beaver, Jared M. Smith, Miki E. Verma, Savannah Norem, Kevin Spakes, Cory Watson, Jeff A. Nichols, Brian Jewell, Michael. D. Iannacone, Chelsey Dunivan Stahl, Kelly M. T. Huffer, T. Sean Oesch

Detecting Industrial Vehicles' Duty Levels Using Contrastive Learning 1630
Luca Cagliero, Silvia Buccafusco, Francesco Vaccarino, Lucia Salvatori, Riccardo Loti

AltOOM: A Data-Driven Out of Memory Root Cause Identification Strategy 1637
Pranjal Chakraborty, Naser Ezzati-Jivan, Vahid Azhari, François Tetreault

A Case Study on ChatGPT Question Generation 1647
Winston Chan, Aijun An, Heidar Davoudi

PersonaSAGE: A Multi-Persona Graph Neural Network 1657
Gautam Choudhary, Iftikhar Ahamath Burhanuddin, Eunyee Koh, Fan Du, Ryan Rossi

Towards Predictive Lifetime-Oriented Temperature Control of Power Electronics in E-Vehicles Via Reinforcement Learning 1667
Anqi Chu, Xiang Xie, Chris Manuel Hermann, Wilhelm Stork, Jörg Roth-Stielow

Ready for the Big Game: How FanDuel Boosts Conversions and Revenue at Scale with a Real-Time Recommender System for Daily Fantasy Sports 1677
Varun Chugh, William Kretz, Steven Rosa, Brooks Beckelman, Hao Tong, Dejan Mitkovski

Applied Data Science for Leasing Score Prediction 1687
Giuseppe Cianci, Roberto Goglia, Riccardo Guidotti, Matteo Kapllaj, Roberto Mosca, Andrea Pugnana, Franco Ricotti, Salvatore Ruggieri

MM-TrafficRisk: A Video-Based Fleet Management Application for Traffic Risk Prediction, Prevention, and Querying 1697
Minh-Son Dao, Muhamad Hilmil Muchtar Aditya Pradana, Koji Zettsu

Near Neighbor Search for Constraint Queries 1707
Gaurav Gupta, Anup Rao, Tung Mai, Ryan A. Rossi, Xiang Chen, Saayan Mitra, Anshumali Shrivastava

Visual Insight Recommendation: From Ranking Insight Visualizations to Insight Types 1716
Camille Harris, Ryan A. Rossi, Sana Malik, Jane Hoffswell, Fan Du, Tak Yeon Lee, Eunyee Koh, Handong Zhao

Load Balancing Algorithms and Their Impacts on Apache Kafka 1726
Hung-Chang Hsiao, Chia-Ping Tsai, Zheng-Xian Li, Chao-Heng Lee, Jia-Sheng Chen, Yu-Chen Lai, Jia-Chi Wang, Shao-Chi Li, Jhih-Cyuan Gao, Yi-Huan Lee

Pricing Used Vehicles at Volkswagen Financial Services AG 1736
Shayan Jawed, Jan Stening, Lars Schmidt-Thieme

A Deep Learning Pipeline for Optimizing Large-Scale Phase Field Simulations 1744
Ramakrishnan Kannan, Cristina Garcia-Cardona, Balasubramaniam Radhakrishnan, Sudip K. Seal

NNCR: Revising Classifications Using Embedding Based K-Nearest-Neighbor Search 1754
Yu-Hsuan Kuo, Saaransh Gulati, Xiaoyu Chu, Pulkit Garg

Uncovering and Understanding Smart TV Users' Picture Quality Preferences Via Big Data Analytics 1764
Hosub Lee

Exploring Approaches to Optimize the Performance of Predictive Coding on Multilanguage Data Sets 1774
Christian J. Mahoney, Nathaniel Huber-Fliflet, Peter Gronvall, Chris Clark, Jianping Zhang, Fusheng Wei, Qiang Mao

Sensorless Monitoring of Shipping Containers 1782
Misaki Mito, Chi Heem Wong, Arnab Chakrabarti

Physical Database Design for Manufacturing Business Analytics 1793
Norifumi Nishikawa, Shinji Fujiwara, Yuto Hayamizu, Kazuo Goda

Will My Flight Be on Time? Learning from Part Failures to Predict Future Reliability 1803
Jongjun Park, Fei Chiang, Eduardo Correia Da Silva, Kevin Lytwyn, Stephen Veldhuis

Revisiting Stylized Facts for Modern Stock Markets 1814
Ethan Ratliff-Crain, Colin M. Van Oort, James Bagrow, Matthew T. K. Koehler, Brian F. Tivnan

Automatic Data Transformation Using Large Language Model - An Experimental Study on Building Energy Data 1824
Ankita Sharma, Xuanmao Li, Hong Guan, Guoxin Sun, Liang Zhang, Lanjun Wang, Kesheng Wu, Lei Cao, Erkang Zhu, Alexander Sim, Teresa Wu, Jia Zou

Improving Conversion Rate Prediction Via Self-Supervised Pre-Training in Online Advertising 1835
Alex Shtoff, Yohay Kaplan, Ariel Raviv

NPRL: Nightly Profile Representation Learning for Early Sepsis Onset Prediction in ICU Trauma Patients 1843
Tucker Stewart, Katherine Stern, Grant O'Keefe, Ankur Teredesai, Juhua Hu

Engineering Data Assets for Public Health Applications: A Covid-19 Case Study 1853
Michael Scriney, Mohan Timilsina, Edward Curry, Lukasz Porwol, Dongyun Nie, Darren Dahley, Jaime B. Fernandez, Mathieu D'Aquin, Mark Roantree

Stock Movement and Volatility Prediction from Tweets, Macroeconomic Factors and Historical Prices 1863
Shengkun Wang, Yangxiao Bai, Taoran Ji, Kaiqun Fu, Linhan Wang, Chang-Tien Lu

Decentralized Personalized Online Federated Learning 1873
Renzhi Wu, Saayan Mitra, Xiang Chen, Anup Rao

Locating Disparities in Machine Learning 1883
Moritz Von Zahn, Oliver Hinz, Stefan Feuerriegel

How Global Retailer ADEO Migrated to Google BigQuery with Database Virtualization 1895
Ehab Abdelhamid, Amirhossein Aleyasen, Michael Duller, Eric Foratier, Vincent Fruleux, Mirella Katch, Gourab Mitra, Rima Mutreja, Jozsef Patvarczki, Matthew Pope, Nikos Tsikoudis, F. Michael Waas

Interpretable Unsupervised Log Anomaly Detection 1899
Jaeho Bang, Sungchul Kim, Ryan Rossi, Tong Yu, Handong Zhao

Fostering Innovation in Urban Transportation Risk Management: A Multi-Sector Collaborative Benchmarking Platform 1903
Minh-Son Dao, Huy Quang Ung, Sadanori Ito, Shinya Wada, Koji Zettsu

Time Series Synthesis Using the Matrix Profile for Anonymization 1908
Audrey Der, Chin-Chia Michael Yeh, Yan Zheng, Junpeng Wang, Huiyuan Chen, Zhongfang Zhuang, Liang Wang, Wei Zhang, Eamonn Keogh

Two-Sample KS Test with approxQuantile in Apache Spark® 1912
Bradley Eck, Duygu Kabakci-Zorlu, Amadou Ba

Theory Meets Practice Approach for Event Correlations 1918
Sai Charan Emmadi, Parag Agrawal, Satya Samudrala, Vikrant Shimpi, Maitreya Natu

Data-Driven Approaches to Detecting Misdeliveries in Truck Logistics Using GPS Data 1922
Ayumu Hidaka, Ryota Shin, Atsushi Tsuchiya, Norihiko Nakabayashi, Yukihiko Okada, Yohei Shida

Explainable Text Classification for Legal Document Review in Construction Delay Disputes 1928
Nathaniel Huber-Fliflet, Jianping Zhang, Peter Gronvall, Fusheng Wei, Philip Spinelli, Adam Dabrowski, Jingchao Yang

Market Size Estimation Model at the Product Level Based on Text Mining Approach 1934
Ye Lim Jung, Hyoung Sun Yoo, Ji Hui Kim

Uncertainty Estimation for a Dual-Embedding Based Entity Extraction Service 1940
Hamed Khorasgani, Namrata Ghadi, Saumil Shah, Henry Zhang

Individualized Reservation Fare Optimization for Designated Driver Service 1945
Kyuyoung Lee, Jiseon Lee, Seungyoun Lee, Hoonyong Shim, Jungmin Kim

Extending the Horizon by Empowering Government Customer Service Officers with ACQAR for Enhanced Citizen Service Delivery 1952
Alvina Hui Shan Lee, Venky Shankararaman, Eng Lieh Ouh

Large Scale Multiplex Latent Graphs for Recommendation in Baidu 1959
Jie Liu, Zhipeng Jin, Yi Yang, Wei Jia, Xuewu Jiao, Shuanglong Li, Lin Liu

Versatile Bengali OCR: Document Analysis Technique for Varied Document Styles and Content 1965
AKM Shahariar Azad Rabby, Hasmot Ali, Md. Majedul Islam, Fuad Rahman

DARC: High-Dimensional Diffusing Anomaly Detection and Root Cause Location in Cloud Computing Systems 1970
Wanning Sun, Xuhua Ma, Ruimin Peng, Yifan Xu, Dongrui Wu

Towards Unified Data Ingestion and Transfer for the Computing Continuum 1978
Muhammad Arslan Tariq, Ovidiu-Cristian Marcu, Grégoire Danoy, Pascal Bouvry

Boosting Feedback: A Framework for Enhancing Ground Truth Data Collection 1982
Kota Tsubouchi, Shuji Yamaguchi, Tatsuru Higurashi

Cost-Efficient Cross-Region Copy in Cloud Data Warehouse: An Evolutionary Approach 1990
Huijun Wu, Maosong Fu, Mansi Kapoor, Zheng Li, Donggeon Kim, Jihyeon Ha, Sungwook Ethan Byun, Prashanth Rajendran, San Tan

Temporal Treasure Hunt: Content-Based Time Series Retrieval System for Discovering Insights..... 1994
Chin-Chia Michael Yeh, Huiyuan Chen, Xin Dai, Yan Zheng, Yujie Fan, Vivian Lai, Junpeng Wang, Audrey Der, Zhongfang Zhuang, Liang Wang, Wei Zhang

Online Forecasting Based Anomaly Detection for Monitoring Large Scale Streaming Data..... 1998
Wei Zhang, David Arbour

Specimen Outlining: A Computational Archival Science Approach 2004
David E. Breen, Andrew Senin, Ajani Levere, Joel Pepper, Jane Greenberg

Accelerating Precision Research and Resolution Through Computational Archival Science Pedagogy 2010
Sarah A. Buchanan, Jennifer L. Wachtel, Jennifer A. Stevenson

AI-Generated Images as an Emergent Record Format 2020
Jessica Bushey

Critical Community-Centeredness: Ethical Considerations for Computational Archival Studies..... 2032
Madelynn Dickerson, Audra Eagle Yun

Authenticating Citizen Journalism Videos by Incorporating the View of Archival Diplomatics into the Verification Processes of Open-Source Investigations (OSINT) 2036
Hoda Amal Hamouda

The Utility of Standards and Good Practice Guidelines for Records Professionals: Comparing Apples, Oranges, and Other Fruits 2047
Shadrack Katuu

Who's in My Archive? an End-To-End Framework for Automatic Annotation of TV Personalities..... 2061
Maurizio Montagnuolo, Fulvio Negro, Alberto Messina, Angelo Bruccoleri, Roberto Iacoviello

An Intelligent Class – the Sequel: The Development of a Novel Context Capturing Method for the Functional Auto Classification of Records 2071
Nathaniel Payne

Analogous Analogues: Digital Twins and Hardware Tracking in GLAM Collections..... 2083
Dian Marie Ross, Edmond Cretu, Victoria Lemieux

Will Blockchain Technology Change How Well National Archives Preserve the Trustworthiness of Digital Records?: Preliminary Results of a Survey 2095
Özhan Saglik, Victoria Lemieux

Can GPT-4 Think Computationally About Digital Archival Practices? 2105
William Underwood, Joan Gage

Exploring the Application of Large Language Models in Detecting and Protecting Personally Identifiable Information in Archival Data: A Comprehensive Study*.....2116
Jianliang Yang, Xiya Zhang, Kai Liang, Yuenan Liu

A Store Evaluation System Using Automatic Scoring of Retail Stores Based on Product Review Analysis 2124
Da Li, Hiroto Nishikawa, Mayumi Ueda, Shinsuke Nakajima

User Latent Interest Estimation in Real Space: A Comparative Analysis of Time-Series and Non-Time-Series Processing Algorithms 2131
Takanobu Omura, Da Li, Panote Siriaraya, Katsumi Tanaka, Yukiko Kawai, Shinsuke Nakajima

Predicting Potential School Shooters from Social Media Posts 2139
Alana Cedeno, Rachel Liang, Sheikh Rabiul Islam

Multimodal One-Class Learning for Malicious Online Content Detection 2146
Roberto Corizzo, Nora Lewis, Lucas P. Damasceno, Allison Shafer, Charles C. Cavalcante, Zois Boukouvalas

Late Fusion-Based Distributed Multimodal Learning 2152
Flavio Giobergia, Elena Baralis

Impact of Mixed Multimodalities and Size Dependence on Performance of Object Detection on Multimodal Satellite Imagery* 2157
Nikita Gordienko, Yuri Gordienko, Oleksandr Rokovyi, Oleg Alienin, Andrii Polukhin, Sergii Stirenko

LoRA-Like Calibration for Multimodal Deception Detection Using ATSFace Data 2163
Shun-Wen Hsiao, Cheng-Yuan Sun

Exploiting Multiple Sequence Lengths in Fast End to End Training for Image Captioning 2173
Jia Cheng Hu, Roberto Cavicchioli, Alessandro Capotondi

Audio-Visual Neural Face Generation with Emotional Stimuli 2183
Juheon Hwang, Jiwoo Kang

Understanding the Language of ADHD and Autism Communities on Social Media 2188
Niloofar Kalantari, Amirreza Payandeh, Marcos Zampieri, Vivian G Motti

Neural Crystals 2196
Sofia Karamintziou, Thanassis Mavropoulos, Dimos Ntioudis, Georgios Meditskos, Stefanos Vrochidis, Ioannis Yiannis Kompatsiaris

Semantic Prompt Based Multi-Scale Transformer for Few-Shot Classification 2200
Hongwu Liu, Shouhong Wan, Peiquan Jin, Xin Wang

Using the CARLA Simulator to Train a Deep Q Self-Driving Car to Control a Real-World Counterpart on a College Campus 2206
Joseph May, Samir Poudel, Sammi Hamdan, Khem Poudel, Jorge Vargas

Enhancing Scientific Image Classification Through Multimodal Learning: Insights from Chest X-Ray and Atomic Force Microscopy Datasets 2211
David C. Meshnick, Nahal Shahini, Debargha Ganguly, Yinghui Wu, Roger H. French, Vipin Chaudhary

De-SaTE: Denoising Self-Attention Transformer Encoders for Li-Ion Battery Health Prognostics 2221
Gaurav Shinde, Rohan Mohapatra, Pooja Krishan, Saptarshi Sengupta

Gender Classification Accuracy Via Two-Dimensional Body Joints Using Convolutional Neural Networks 2229
Cheng-En Sung, Nada Attar

New Finger Photo Databases with Presentation Attacks and Demographics 2234
Anudeep Vurity, Emanuela Marasco

CLIP-PubOp: A CLIP-Based Multimodal Representation Fusion Method for Public Opinion 2243
Zhibo Wang, Yi Guo, Jiaojiao Fu

Multimodal Large Language Models: A Survey 2247
Jiayang Wu, Wensheng Gan, Zefeng Chen, Shicheng Wan, Philip S. Yu

Evaluating CLIP's Understanding on Relationships in a Blocks World 2257
Kairui Zhang, Martha Lewis

MPI4Spark Meets YARN: Enhancing MPI4Spark Through YARN Support for HPC 2265
Kinan Al-Attar, Aamir Shafi, Hari Subramoni, Dhabaleswar K. Panda

Optimized Model of Ledger Database Management to Handle Vehicle Registration 2275
Neha Gaonkar, Parnia Samimi, Jagdev Bhogal, Luke Porter, Rob Squire

BFRecSys: A Blockchain-Based Federated Matrix Factorization for Recommendation Systems 2283
Dongkun Hou, Jie Zhang

Deep Reinforcement Learning Based Data Placement Optimization in Data Center Networks 2293
Ravi Kaler, Durga Toshniwal

Blockchain-Based Verifiable and Reliable File Access Control Layer for Cloud Storages 2303
Daun Kim, Hyunjoo Yang, Howon Lee, Sejin Park

Lazy Read: Asynchronous Execution of Synchronous File I/O 2311
Ansgar Lößer, Florian Raskob, Björn Scheuermann

Optimal Repair Algorithm of Single-Disk Failure for Array Codes with Local Properties 2319
Chuang Luo, Hanxu Hou, Peng Zhang, Hui Li

H-MIS: A Hierarchical Multi-Identifier System Based on Blockchain 2326
Qi Lyu, Hui Li, Xinnan Lin, Han Wang, Hanxu Hou, Yuguo Yin, Qianbin Chen, Ping Fan, Gang Wan, Xiang Zhu, Selwyn Deng, Weixiang Huang, Jianping Wu, Jieren Cheng, Wenyuan Yang

Design and Implementation of an ERC-20 Smart Contract on the Ethereum Blockchain 2334
Joshua Priest, Cameron Cooper, Savvy Lovell, Yong Shi, Dan Lo

Towards a Peer-To-Peer Data Distribution Layer for Efficient and Collaborative Resource Optimization of Distributed Dataflow Applications 2339
Dominik Scheinert, Soeren Becker, Jonathan Will, Luis Englaender, Lauritz Thamsen

RNBFT: Leveraging Randomness to Achieve Scalable Byzantine Consensus 2346
Parth Shukla, Saeed Samet

A Federated Learning Scheme Based on Lightweight Differential Privacy 2356
Wenlong Song, Hong Chen, Zhijie Qiu, Lei Luo

Reed-Solomon Codes Over Ring with Lower Computational Complexity 2362
Shicheng Tan, Hanxu Hou, Leilei Yu

SmartDT: An Effective Vulnerability Detection System of Smart Contracts Based on Deep Learning 2369
Xiaozhou You, Hui Li, Han Wang, Faisal Mehmood

Data-Push Projects and Their Unique Feature: Managing with Anomalies 2377
Antoine Bordas, Pascal Le Masson, Benoit Weil

Big Data Tools: Interoperability Study and Performance Testing 2386
Asma Dhaouadi, William Paccoud, Khadija Bousselmi, Sébastien Monnet, Mohamed Mohsen Gammoudi, Slimane Hammoudi

MLOps in Data Science Projects: A Review 2396
Christian Haertel, Daniel Staegemann, Christian Daase, Matthias Pohl, Abdulrahman Nahhas, Klaus Turowski

A Capacity Framework of Community Readiness for Supporting Big Data Science Projects During Cyberinfrastructure Diffusion..... 2405
Kerk F. Kee, Alex Olshansky, Shan Xu, Kulsawasd Jitkajornwanich

ExplainerX: An Integrated and Explainable AI Framework for Nearly Zero-Energy Buildings.....2411
Asma Kermiche, Chaima Rouzzi, Ladjel Bellatreche

Bridging the Gap in AI-Driven Workflows: The Case for Domain-Specific Generative Bots 2421
Akit Kumar, M. S. Lakshmi Devi, Jeffrey S. Saltz

Data Science Failure: A Literature Review 2431
Sucheta Lahiri, Jeff Saltz

Extending Variability-Aware Model Selection with Bias Detection in Machine Learning Projects..... 2441
Cristina Tavares, Nathalia Nascimento, Paulo Alencar, Donald Cowan

Understanding Data Pre-Processing with a Hybrid Diagram Integrating ER and Data Flow Notation..... 2450
Robin Varghese, Carlos Ordonez

Predictive Analytics in IoT and CPS: Enhancing Industrial Machinery Reliability Through Sensor Data-Driven Remaining Useful Life Estimation 2456
Emrullah Gultekin, Mehmet S. Aktas

Dynamic Analysis for the Detection of Locked Ether Smart Contracts 2466
Zulfiqar Ali Khan, Akbar Siami Namin

Interaction in Metaverse: A Survey 2473
Hong Lin, Zirun Gan, Wensheng Gan, Zhenlian Qi, Yuehua Wang, Philip S. Yu

Detecting Phishing URLs Using the BERT Transformer Model..... 2483
Denish Omondi Otieno, Faranak Abri, Akbar Siami Namin, Keith S. Jones

AGV Quality of Service Throughput Prediction Via Neural Networks..... 2493
Katarzyna Prokop, Dawid Polap, Gautam Srivastava

The Performance of Machine and Deep Learning Algorithms in Detecting Fake Reviews 2499
Bharkavi Sachithanandam, Akbar Siami Namin, Faranak Abri

Exploiting Large Language Models (LLMs) Through Deception Techniques and Persuasion Principles..... 2508
Sonali Singh, Faranak Abri, Akbar Siami Namin

An Iterative Sinogram Metal Artifact Reducdion Based on UNet..... 2518
Zichong An, Xuemei Zhu, Xiangling Fu, Junqi Ma, Chenyi Guo, Zheng Zhang

Learning the Dynamics of Concentration Fields in Vascular Stenosis with Deep Hidden Physics Models..... 2523
Rebecca Kador, Helen Schneider, David Biesner, Babette Dellen, Rafet Sifa

Deployment and Application of Deep Learning Models Under Computational Constraints..... 2529
Jinwen Li, Wendong Wang

A Pose and Shape-Aware Cross-Skeleton Motion Retargeting Framework..... 2534
Chengyu Liu, Bo Zhang, Wendong Wang

GPU-Accelerated PostgreSQL for Scalable Management and Processing of Irregular Time-Series Data Using SPI..... 2541
Jamal Saied-Walker, Pallavi Gupta, Ruhan Yi, Noah Marchal, Marjorie Skubic, Grant J. Scott

Learning to Generate Radiology Findings from Impressions Based on Large Language Model..... 2550
Weilong Wu, Miao Li, Ji Wu, Ming Ni, Huishu Yuan

A 3D Framework for Brain Tumor Segmentation from Multi-Modal MR Images..... 2555
Shuo Yan, Zheng Zhang, Bo Zhang, Wendong Wang

Optimization of Dangerous Goods Stowage: A Hybrid Approach Between Faster R-CNN and Coronavirus Metaheuristic..... 2561
Chaabane Abdelali, Yachba Khadidja, Ladjel Bellatreche

Proposal of an AI Based Approach for Urban Traffic Prediction from Mobility Data..... 2570
Mariaelena Berlotti, Sarah Di Grande, Salvatore Cavalieri, Vincenza Torrisi, Giuseppe Inturri

Driving into Uncertainty: An Adversarial Generative Approach for Multivariate Scenario Generation..... 2578
Michele Carbonera, Michele Ciavotta, Enza Messina

Charting the Course of Ship Track Prediction: A Novel Approach for Maritime Traffic Analysis and Enhanced Situational Awareness..... 2588
Lubna Eljabu, Mohammad Etemad, Stan Matwin

ChattSpeed: Toward a New Dataset for Single Camera Visual Speed Estimation for Urban Testbeds..... 2598
Yasir Hassan, Mina Sartipi

Leveraging Probe Data and Machine Learning to Derive and Interpret Macroscopic Fundamental Diagrams Across U.S. Cities..... 2606
Ling Jin, Xiaodan Xu, Yuhan Wang, Kaveh Farokhi Sadabadi, Alina Lazar, Duleep Rathgamage Don, Zachary Needell, C Anna Spurlock, Mahyar Amirgholy, Mona Asudegi

Empowering Search and Rescue Operations with Big Data Technology: A Comprehensive Study of YOLOv8 Transfer Learning for Transportation Safety..... 2616
Luciano Lima, Fabio Andrade, Youcef Djenouri, Carlos Pfeiffer, Marcos Moura

Unstructured Transportation Safety Board Findings Categorization Using the Knowledge Graph Pipeline..... 2624
Ritesh Panday, Chung-Horng Lung

Identifying Regions of High Demand for Transportation Services Based on Cluster Evolution and Graph Analysis 2633
Ivens Portugal, Paulo Alencar, Donald Cowan

MixPCD: A Novel Method of Constructing Driving Cycle for Electric Bus with Sparse Data 2639
Xingxing Wang, Peilin Ye, Yelin Deng, Yinnan Yuan, Yu Zhu, Hongjun Ni

Enhancing Concept Drift Detection in Drifting and Imbalanced Data Streams Through Meta-Learning 2648
Gabriel Jonas Aguiar, Alberto Cano

AutoClass: AutoML for Data Stream Classification..... 2658
Maroua Bahri, Nikolaos Georgantas

Toward Parallelization Technique for Stream-Based Lossless Data Compression..... 2667
Taiki Kato, Shinichi Yamagiwa, Koichi Wada

Continuous Soccer Pass Detection: A Comparison Between Traditional and Streaming Machine Learning Methods..... 2673
Stefania Menconi, Giacomo Ziffer, Alessio Bernardo, Emanuele Della Valle

Online Closed Episode Mining with Root-Order Decomposition..... 2683
Hinano Sako, Yoshitaka Yamamoto

Mitigating Concept Drift in Distributed Contexts with Dynamic Repository of Federated Models 2690
Elena Tsiporkova, Michiel De Vis, Sarah Klein, Anna Hristoskova, Veselka Boeva

Trustworthy Medical Operational AI: Marrying AI and Regulatory Requirements 2700
Fabian Berns, Georg Zimmermann, Christian Borgelt, Niclas Heilig, Jan Kirchhoff, Florian Stumpe

Leveraging Deep Learning to Improve COVID-19 Forecasting Using Wastewater Viral Load 2705
Mehrdad Fazli, Heman Shakeri

Leveraging Big Data from Smart Thermostats: A Vision to Evaluate Heatwave Impacts on Sleep Health in the Elderly [Vision Paper]..... 2714
Jasleen Kaur, Vivek Chauhan, Arlene Oetomo, Kang Wang, Plinio Morita

Hyperparameter Tuning of a Deep Learning EEG-Based Neural Network for the Diagnosis of ADHD 2719
Javier Sanchis, Miguel A. Teruel, Juan Trujillo

Method for Creating Privacy-Preserving Information Using Probabilistic Latent Semantic Analysis..... 2726
Yuki Sugawara, Eiichi Sakurai, Yoichi Motomura, Yukihiko Okada, Kai Tanabe, Akiko Tsukao, Shinya Kuno

Extraction of Medication and Temporal Relation from Clinical Text Using Neural Language Models..... 2735
Hangyu Tu, Lifeng Han, Goran Nenadic

Developing and Evaluating a Model-Based Metric for Legal Question Answering Systems 2745
Dilan Bakir, Beytullah Yildiz, Mehmet S. Aktas

Exploring the Performance Impacts of Training Predictive Models with Inclusive Email Threads..... 2755
Chris Clark, Han Qin, Nathaniel Huber-Fliflet, Adam Dabrowski, Jianping Zhang

Explainable e-Discovery (XeD) Using an Interpretable Fuzzy ARTMAP Neural Network for Technology-Assisted Review 2761
Charles Courchaine, Tasnova Tabassum, Corey Wade, Ricky J. Sethi

Towards Open-Set Contract Clause Recognition 2767
Angelo Impedovo, Giuseppe Rizzo, Antonio Di Mauro

Graph-Based Approach for European Law Classification 2777
Raffaele Russo, Giuliano Di Giuseppe, Alessandro Vanacore, Valerio La Gatta, Antonino Ferraro, Antonio Galli, Marco Postiglione, Vincenzo Moscato

Empirical Study of LLM Fine-Tuning for Text Classification in Legal Document Review 2786
Fusheng Wei, Robert Keeling, Nathaniel Huber-Fliflet, Jianping Zhang, Adam Dabrowski, Jingchao Yang, Qiang Mao, Han Qin

An Empirical Analysis of Text Segmentation for BERT Classification in Extended Documents 2793
Jingchao Yang, Fusheng Wei, Nathaniel Huber-Fliflet, Adam Dabrowski, Qiang Mao, Han Qin

BBVA AI Factory at FinCausal 2023: A RoBERTa Fine-Tuned Model for Causal Detection 2798
Alberto Algarra, David Muelas

Advanced NLP Techniques for Summarizing Multilingual Financial Narratives from Global Annual Reports 2802
Dilshod Azizov, Jiyong Li, Hilal Alquabeh, Shangsong Liang

Identifying Going Concern Issues in Auditor Opinions: Link to Bankruptcy Events 2805
Konstantinos Bougiatiotis, Elias Zavitsanos, Georgios Paliouras

Uncovering Inconsistencies and Contradictions in Financial Reports Using Large Language Models 2814
Tobias Deußer, David Leonhard, Lars Hillebrand, Armin Berger, Mohamed Khaled, Sarah Heiden, Tim Dilmaghani, Bernd Kliem, Rüdiger Loitz, Christian Bauckhage, Rafet Sifa

Automatic ESG Assessment of Companies by Mining and Evaluating Media Coverage Data: NLP Approach and Tool 2823
Jannik Fischbach, Max Adam, Victor Dzhagatspanyan, Daniel Mendez, Julian Frattini, Oleksandr Kosenkov, Parisa Elahidoost

Humans Vs. ChatGPT: Evaluating Annotation Methods for Financial Corpora 2831
Jamshed Kaikaus, Haoen Li, Robert J. Brunner

Advancements in Financial Document Structure Extraction: Insights from Five Years of FinTOC (2019-2023) 2839
Juyeon Kang, Mauli Mehulkumar Patel, Anushka Agrawal, Simhadri Sevitha, R Srinivasa, Sandra Bellato, M Anand Kumar, Ngawang Dempa Tsang, Mo El-Haj

Impact Analysis of Social Events on Industries Through Narrative Causal Search 2845
Ryotaro Kobayashi, Yuri Murayama, Kiyoshi Izumi

The Financial Document Causality Detection Shared Task (FinCausal 2023) 2855
Antonio Moreno-Sandoval, Jordi Porta-Zamorano, Blanca Carbajo-Coronado, Doaa Samy, Dominique Mariko, Mahmoud El-Haj

Topic-Sentiment Analysis of Central Bank Press Conferences: BOJ Case Study 2861
Kazuhiro Seki, Masahiko Shibamoto, Takashi Kamihigashi

Investigating Large Language Models for Financial Causality Detection in Multilingual Setup 2866
Neelesh K Shukla, Raghu Katikeri, Msp Raja, Gowtham Sivam, Shlok Yadav, Amit Vaid, Shreenivas Prabhakararao

Generative AI Approach to Distributed Summarization of Financial Narratives 2872
Neelesh K Shukla, Raghu Katikeri, Msp Raja, Gowtham Sivam, Shlok Yadav, Amit Vaid, Shreenivas Prabhakararao

Summarizing Financial Reports with Positional Language Model 2877
Natalia Vanetik, Elizaveta Podkaminer, Marina Litvak

Meticulously Analyzing ESG Disclosure: A Data-Driven Approach 2884
Tik Yu Yim, Yuxuan Zhang, Wenting Tan, Tak-Wah Lam, Siu Ming Yiu

The Financial Narrative Summarisation Shared Task (FNS 2023) 2890
Elias Zavitsanos, Aris Kosmopoulos, George Giannakopoulos, Marina Litvak, Blanca Carbajo-Coronado, Antonio Moreno-Sandoval, Mo El-Haj

IoT-Reg: A Comprehensive Knowledge Graph for Real-Time IoT Data Privacy Compliance 2897
Kelvin Uzoma Echenim, Karuna Pande Joshi

Characterizing Evolutionary Trends in Temporal Knowledge Graphs with Linear Temporal Logic 2907
Valeria Fionda, Giuseppe Pirró

Leveraging Knowledge Graphs for Matching Heterogeneous Entities and Explanation 2910
Sahar Ghassabi, Behshid Behkamal, Mostafa Milani

STROOBnet Optimization Via GPU-Accelerated Proximal Recurrence Strategies 2920
Ted Edward Holmberg, Mahdi Abdelguerfi, Elias Ioup

GCN-Based Explainable Recommendation Using a Knowledge Graph and a Language Model 2930
Jeongbin Lee, Kunyoung Kim, Mye Sohn, Jongmo Kim

Ontology-Based Generation of Data Platform Assets 2937
Vincenzo De Leo, Gianni Fenu, David Greco, Nicolò Bidotti, Paolo Platter, Enrico Motta, Andrea Giovanni Nuzzolese, Francesco Osborne, Diego Reforgiato Recupero

Knowledge Graphs for Competency-Based Education 2942
Andrea Linxen, Florian Endel, Simone Opel, Christian Beecks

Force-Directed Graph Embedding with Hops Distance 2946
Hamidreza Lotfalizadeh, Mohammad Al Hasan

Knowledge Graphs in Spatial-Temporal Cluster Evolution Analysis 2954
Ivens Portugal, Paulo Alencar, Donald Cowan

Supporting Practical URI Mappings in Virtual Knowledge Graph-Based Relational Data Integration 2958
Shogo Sato, Tadashi Masuda, Toshiyuki Amagasa

BUILD-KG: Integrating Heterogeneous Data into Analytics-Enabling Knowledge Graphs 2965
Kara Schatz, Pei-Yu Hou, Alexey V. Gulyuk, Yaroslava G. Yingling, Rada Chirkova

Knowledge Prompt for Whisper: An ASR Entity Correction Approach with Knowledge Base 2975
Min Zhang, Xiaosong Qiao, Yanqing Zhao, Chang Su, Yinglu Li, Yuang Li, Ming Zhu, Mengyao Piao, Song Peng, Shimin Tao, Hao Yang, Yanfei Jiang

CERBERE: Cybersecurity Exercise for Red and Blue Team Entertainment, REproducibility 2980
Pierre-Victor Besson, Romain Brisse, Hélène Orsini, Natan Talon, Jean-François Lalande, Frédéric Majorczyk, Alexandre Sanchez, Valérie Viet Triem Tong

Hide My Payload: An Empirical Study of Antimalware Evasion Tools 2989
Hannes Holm, Erik Hyllienmark

Automated Cyber Threat Intelligence Generation on Multi-Host Network Incidents 2999
Cristoffer Leite, Jerry Den Hartog, Daniel R. Dos Santos, Elisa Costante

AMINet: An Industrial Honeynet for AMI Systems 3009
Alexios Lekidis, George Daniil, Panagiotis Mavrommatis, Konstantinos Lampropoulos, Odysseas Koufopavlou

Evaluating Representativeness in PDF Malware Datasets: A Comparative Study and a New Dataset 3017
Ran Liu, Robert Joyce, Cynthia Matuszek, Charles Nicholas

ICS Honeypot Interactions: A Latitudinal Study 3025
Francesco Lupia, Marco Lucchese, Massimo Merro, Nicola Zannone

Enhancing Cyber Threat Hunting: A Visual Approach with the Forensic Visualization Toolkit 3035
Jihane Najar, Marinos Tsantekidis, Aris Sotiropoulos, Vassilis Prevelakis

Machine Learning–Based Security Alert Screening with Focal Loss 3043
Samuel Ndichu, Tao Ban, Takeshi Takahashi, Daisuke Inoue

AGIR: Automating Cyber Threat Intelligence Reporting with Natural Language Generation 3053
Filippo Perrina, Francesco Marchiori, Mauro Conti, Nino Vincenzo Verde

CVE Representation to Build Attack Positions Graphs 3063
Manuel Poisson, Valérie Viet Triem Tong, Gilles Guette, Frédéric Guihéry, Damien Crémilleux

Investigating Initial Access Strategies and Malware Deployment Tactics Used by Iranian Advanced Persistent Threat Groups 3068
Ryan Simons

Assessing the Threat Level of Software Supply Chains with the Log Model 3079
Luís Soeiro, Thomas Robert, Stefano Zacchiroli

A Modular Approach to Automatic Cyber Threat Attribution Using Opinion Pools 3089
Koen T. W. Teuwen

Learning Explainable Multi-View Representations for Malware Authorship Attribution 3099
Irsyad Adam, Alex Waagen, Dana Warmsley, Jiejun Xu

Hate Speech and Hate Crimes: A Data-Driven Study of Evolving Discourse Around Marginalized Groups 3107
Malvina Bozhidarova, Jonathn Chang, Aaishah Ale-Rasool, Yuxiang Liu, Chongyao Ma, Andrea L. Bertozzi, P. Jeffrey Brantingham, Junyuan Lin, Sanjukta Krishnagopal

AutoKG: Efficient Automated Knowledge Graph Generation for Language Models 3117
Bohan Chen, Andrea L. Bertozzi

Hybrid Attack Graph Generation with Graph Convolutional Deep-Q Learning 3127
Sam Donald, Rounak Meyur, Sumit Purohit

A Multi-Source Graph Database to Showcase a Recommender System for Dyslexic Students..... 3134
Karim El Hage, Adel Remadi, Yasmina Hobeika, Ruining Ma, Victor Hong, Francesca Bugiotti

Geospatial Enrichment of Urban Data for Advanced City Planning: A Pilot Study 3139
Iva Krasteva, Dessislava Petrova-Antonova, Flavio De Paoli, Emil Hristov, Milena Borukova, Michele Ciavotta, Roberto Avogadro

LLMs with User-Defined Prompts as Generic Data Operators for Reliable Data Processing..... 3144
Luyi Ma, Nikhil Thakurdesai, Jiao Chen, Jianpeng Xu, Evren Korpeoglu, Sushant Kumar, Kannan Achan

Combining Fault-Tolerant Persistence and Low-Latency Streaming Access to Binary Data for AI Models 3149
Gabriele Scaffidi Militone, Daniele Apiletti, Giovanni Malnati

Fast Meta Failure Recovery for Federated Meta-Learning..... 3154
Brandon Delliquadri, Chao Wang, Shuo Chen, Zhengxiong Li, Hui Luo, Hailu Xu

Digital Finance, Market Competition, and Risk-Taking: A Study of Rural Commercial Banks in China 3159
Chongshuang Hu, Minkang Li, Hufeng Yang, Minghao Li, Jichao Li, Jiang Jiang

Toward a Novel Methodology in Economic Experiments: Simulation of the Ultimatum Game with Large Language Models 3168
Ayato Kitadai, Yudai Tsurusaki, Yusuke Fukasawa, Nariaki Nishino

CultureBERT: Measuring Corporate Culture with Transformer-Based Language Models 3176
Sebastian Koch, Stefan Pasch

Lower Alertness Entails More Cooperation: Evidence from Prisoner's Dilemma and Coordination Games..... 3185
Toshio Kokubo, Nobuhito Kon, Morimitsu Kurino, Wakuo Saito, Natsumi Shimada, Masashi Yanagisawa

Does Double-Entry Bookkeeping Information Generated Using Node2vec Contribute to Forecasting Future Performance?..... 3194
Ryoki Motai, Masato Kamebuchi, Sora Watanabe, Ryo Matsumoto, Yukihiko Okada

Conducting an Experiment at Multiple Sites with Small Subject Pools: How is Raven Score Effective as a Covariate? 3202
Kazuhito Ogawa, Yusuke Osaki, Tetsuya Kawamura, Hiromasa Takahashi, Satoshi Taguchi, Yoichiro Fujii, Kengo Kurosaka, Kohei Iyori, Naoki Watanabe

Port Arrival Reservation System Using Auctions for Fuel Consumption Reduction 3212
Rintaro Rai, Takuya Nakashima, Nariaki Nishino

Mathematical Model Analysis of Photovoltaic-Induced Land Use Transition 3222
Yumi Watahiki, Hiroyasu Ando, Isamu Takahara

Efficiency in a College Admission Experiment: A Preliminary Simulation with Cognitive Ability Scores 3231
Naoki Watanabe, Tetsuya Kawamura, Kazuhito Ogawa

Evaluating the Alleviation of Traffic Congestion Through Bus Rapid Transit Using Multi-Agent Simulation 3237
Zeyu Zhou, Mizuki Kobayashi, Uta Sato, Kazuma Akashi, Ayato Kitadai, Soma Sugihara, Yusuke Fukasawa, Masanori Fujuta, Nariaki Nishino

Improving Fingerprint Classification Accuracy Using Image Enhancement Deep Learning Models 3245
Nikita Dalvi, Min Kyung An, ABM Rezbaul Islam, Vung Pham

Distributed Continual Intrusion Detection: A Collaborative Replay Framework 3255
Kamil Faber, Bartlomiej Sniezynski, Roberto Corizzo

ODTT: Optimized Dynamic Taxonomy Tree with Differential Privacy 3264
Yijie Gui, Qizhi Chen, Wensheng Gan, Yongdong Wu

Pruning Federated Learning Models for Anomaly Detection in Resource-Constrained Environments 3274
Simone Magnani, Stefano Braghin, Ambrish Rawat, Roberto Doriguzzi-Corin, Mark Purcell, Domenico Siracusa

IoT Botnet-Traffic Classification Using Few-Shot Learning 3284
Davide Di Monda, Giampaolo Bovenzi, Antonio Montieri, Valerio Persico, Antonio Pescapè

EEG Signal-Based Authentication: A Performance Evaluation of Feature Extraction and Classification Techniques 3294
Rahul Nagarajan, Malemsana Thokchom, Abdullah Irfan Siddiqui, Mohammad I. Husain

Hierarchical Federated Transfer Learning and Digital Twin Enhanced Secure Cooperative Smart Farming 3304
Lopamudra Praharaj, Maanak Gupta, Deepti Gupta

Analyzing Data Sets for ML-Driven Fraud Detection in SAP Systems 3314
Jonas Schnepf, Bernd Scheuermann, Paula Vetter

An Experimental Study on Random Projection Based Biometric Security 3325
Rishabh Shukla, Harkeerat Kaur

A Graph Data Model-Based Micro-Provenance Approach for Multi-Level Provenance Exploration in End-To-End Climate Workflows 3332
Sandro Fiore, Mattia Rampazzo, Donatello Elia, Ludovica Sacco, Fabrizio Antonio, Paola Nassisi

Evaluate and Compare Machine Learning Models for Temperature Forecasting of Meteorological Data in Hohhot 3340
Ruyuan Han, Lei Yu

TensorBank: Tensor Lakehouse for Foundation Model Training 3350
Romeo Kienzler, Johannes Schmude, Naomi Simumba, Benedikt Blumenstiel, Marcus Freitag, Daiki Kimura, Zoltan Arnold Nagy, Michael Behrendt, Hendrik Hamann, S. Karthik Mukkavilli, Daniel Salles Civitarese

Proximal Policy Optimization Based Computations Offloading for Delay Optimization in UAV-Assisted Mobile Edge Computing 3355
Praveen Kumar, Priyadarshni, Shivani Tripathi, Rajiv Misra

Interactive Visualization and Portable Image Blending of Massive Aerial Image Mosaics 3365
Steve Petruzza, Brian Summa, Amy Gooch, Christine M. Laney, Tristan Goulden, John Schreiner, Steven Callahan, Valerio Pascucci

An Early Warning System for Air Pollution Surveillance: A Big Data Framework to Monitoring Risks Associated with Air Pollution 3371
Shahan Salim, Irfhana Zakir Hussain, Jasleen Kaur, Plinio P. Morita

Estimating Water Storage from Images 3375
Ajmal Shahbaz, Syed Yunas, Lyndon Smith, Chad Staddon

Time Series Transformer for Long Term Rainfall Forecasting Towards Water Distribution Management in Smart Cities. 3380
Aswini Ghosh

Workload Shifting Based on Low Carbon Intensity Periods: A Framework for Reducing Carbon Emissions in Cloud Computing 3387
Shivani Tripathi, Praveen Kumar, Priyadarshni Gupta, Rajiv Misra, T. N. Singh

Analysis and Forecasting of Carbon Emission in SAARC Countries Using Attention-Based LSTM 3396
Anil Verma, Harshit Dhankhar, Rajiv Misra, T. N. Singh, Om Prakash Dhakal

SRNI-CAR: A Comprehensive Dataset for Analyzing the Chinese Automotive Market 3405
Ruixin Ding, Bowei Chen, James M. Wilson, Zhi Yan, Yufei Huang

Forecasting Commodities Prices with the Bayesian Symbolic Regression Compared to Other Methods 3413
Krzysztof Drachal

Crude Oil Price Forecasting Base on an EEMD and Multi-Scale Time Series Analysis Combination Methods 3422
Ke Huang, Yin Luo, Lifei Ke, Jianwen Qin, Pingping Liu

DWT-TimesNet: An Improved Model for Wheat Price Long-Time Series Forecasting 3431
Mingwei Lei, Xingfen Wang

Dynamic Price Personalization: Predicting the Optimum Price to Maximize Revenue 3439
Kazuhiro Miyatsu

GDP Nowcasting Based on Daily Electricity Data and Financial Market Data 3445
Wei Shang, Xun Zhang, Rui Tao, Xiuzhenzi Wang, Lili Zhang

Colocation Datacenter Customer Power Usage Forecasting Using Synthetic Data and Integration of Macroeconomic Indicators 3453
Neda Zarayeneh, Malarvizhi Sankaranarayanasamy, Pegah Mavaie, Omanshu Thapliyal, Prasun Singh, Ravigopal Vennelakanti

Artificial Intelligence-Based Plant Breeding Using Genotype and Phenotype Data: Methods and Future Work 3458
Jiho Choi, Sung-Woo Byun, Taehoon Lim, Hye In Lee, Hwa Seon Shin, Geon Woo Kim, Jin-Kyung Kwon, Byoung-Cheorl Kang

Synergizing Measurement Science and Artificial Intelligence in Smart Agriculture 3464
Francesco Lamonaca, Domenico Luca Carnì

Deep Machine Learning on Segmenting and Classifying Crop Images Taken by Unmanned Aerial Vehicle 3470
Dan Chia-Tien Lo, Bobin Deng, Yong Shi

From Satellites to Fields: Machine Learning Applications for Prediction of Corn Production Using NDVI, Precipitation and Land Surface Temperature for Large Producer Countries 3479
Saniya Nangia, Thilanka Munasinghe, Heidi Tubbs, Assaf Anyamba

Transforming Agriculture with Intelligent Data Management and Insights 3489
Yu Pan, Jianxin Sun, Hongfeng Yu, Geng Bai, Yufeng Ge, Joe Luck, Tala Awada

Crop Loss Estimation in Maize Agriculture: A Deep Learning Perspective..... 3499
Geeta Rani, Vijaypal Singh Dhaka, Eugenio Vocaturo, Ester Zumpano

Visualization of 3D Hyperspectral Soil Mapping Data Via Autoencoder-Based Clustering 3503
Jianxin Sun, Xinyan Xie, Yu Pan, Yakub Islamov, Yufeng Ge, Hongfeng Yu

Precision Grazing: When Agriculture, Livestock and Technology Unite 3510
Oscar Tamburis, Mario Magliulo, Adriano Tramontano, Giulio Perillo, Arriel Benis, Serena Calabrò

A Low-Cost IoT-Based Smart Farming System for Crop Recommendation and Resource Management 3516
Daniel Timko, Adityan Elangovan, Aruna Elangovan, Mike Sharko, Ali Ahmadinia

Tracking and Predicting Productions in Agricultural Processes: Applications and Experiences..... 3524
Patrizia Vizza, Giuseppe Timpano, Francesco Vescio, Gianmichele Caligiuri, Fulvia Caligiuri, Pasquale Lambardi, Pierangelo Veltri, Pietro H Guzzi, Giuseppe Tradigo

AI-Driven Agriculture: Opportunities and Challenges 3530
Eugenio Vocaturo, Geeta Rani, Vijaypal Singh Dhaka, Ester Zumpano

Integrating Climate Variable Data in Machine Learning Models for Predictive Analytics of Tomato Yields in California..... 3538
Tianze Zhu, Junyi Wu, Tingyi Tan, Shuheng Wang, Thilanka Munasinghe, Heidi Tubbs, Assaf Anyamba

Adaptive Intrusion Detection Systems: Class Incremental Learning for IoT Emerging Threats..... 3547
Francesco Cerasuolo, Giampaolo Bovenzi, Christian Marescalco, Francesco Cirillo, Domenico Ciuonzo, Antonio Pescapè

One to Rule Them All: A Study on Requirement Management Tools for the Development of Modern AI-Based Software 3556
Abdul-Rasheed Ottun, Mehrdad Asadi, Michell Boerger, Nikolay Tcholtchev, João Gonçalves, Dusan Borovcanin, Bartlomiej Siniarsk, Huber Flores

A Max-Min Security Game for Coordinated Backdoor Attacks on Federated Learning..... 3566
Omar Abdel Wahab, Anderson Avila

A Novel Approach to Parallel Anomaly Detection: Application in Cybersecurity 3574
Ziani Zineb, Emad Nahid, Bouaziz Ahmed

Defending Federated Learning Against Model Poisoning Attacks 3584
Ibraheem Aloran, Saeed Samet

Transformer-Based LLMs in Cybersecurity: An In-Depth Study on Log Anomaly Detection and Conversational Defense Mechanisms..... 3590
Prasasthy Balasubramanian, Justin Seby, Panos Kostakos

Improving Cybersecurity Awareness: Tweet Classification Using Multilingual Sentence Embeddings and Contextual Features 3600
Anastasia Cotov, Carlo Bono, Cinzia Cappiello, Barbara Pernici

A Cascade CNN Model Based on Adaptive Learning Rate Thresholding for Reliable Face Recognition 3608
Imen Labiadh, Larbi Boubchir, Hassene Seddik

Clean-Label Poisoning Attack Against Fake News Detection Models 3614
Jiayi Liang, Xi Zhang, Yuming Shang, Sanchuan Guo, Chaozhuo Li

Exploring the Distinctive Tweeting Patterns of Toxic Twitter Users 3624
Hina Qayyum, Muhammad Ikram, Benjamin Zi Hao Zhao, Ian D. Wood, Nicolas Kourtellis, Mohamed Ali Kaafar

On Mission Twitter Profiles: A Study of Selective Toxic Behavior 3634
Hina Qayyum, Muhammad Ikram, Benjamin Zi Hao Zhao, Ian D. Wood, Nicolas Kourtellis, Mohamed Ali Kaafar

Complementary Attention-Based Deep Learning Detection of Fake Faces 3644
Supriyo Sadhya, Xiaojun Qi

Recent Artificial Intelligence Advances in Detection and Diagnosis of Sickle Cell Disease: A Review 3650
Abdourahmane Balde, Avewe Bassene, Lamine Faty, Mamadou Soumboundou, Ousmane Sall, Youssou Faye

Exploring the Value of Pre-Trained Language Models for Clinical Named Entity Recognition 3660
Samuel Belkadi, Lifeng Han, Yuping Wu, Goran Nenadic

Graph Encoding-Enhanced Transformer for Drug Recommendation 3670
Xunsheng Cai, Syauki Aulia Thamrin, Arbee L. P. Chen

Interpretable and Robust Hospital Readmission Predictions from Electronic Health Records 3679
Hugo Calero-Díaz, Rebeen Ali Hamad, Christian Atallah, John Casement, Dexter Canoy, Nick J Reynolds, Michael Barnes, Paolo Missier

How Much Data Do I Need? a Case Study on Medical Data 3688
Ayse Betul Cengiz, A. Stephen McGough

Data Poisoning Attacks Over Diabetic Retinopathy Images Classification 3698
Fabio Martinelli, Francesco Mercaldo, Marcello Di Giammarco, Antonella Santone

NCL-SM: A Fully Annotated Dataset of Images from Human Skeletal Muscle Biopsies 3704
Atif Khan, Conor Lawless, Amy E Vincent, Charlotte Warren, Valeria Di Leo, Tiago Gomes, A. Stephen McGough

A Novel Ensemble-Based Framework for Feature Identification and Classification of COVID-19 Electronic Health Record Data 3711
Asmita Mahajan, Durga Toshniwal

Forecasting COVID-19 Hotspots in Florida Public Schools: A Machine Learning Approach 3721
Mingming Peng, Askal Ayalew Ali, Hongmei Chi

Enhancing Vocal-Based Laryngeal Cancer Screening with Additional Patient Information and Voice Signal Embedding 3731
Jaemin Song, Yong Oh Lee, Seho Park, Youn Kyu Lee, Hansang Park, Hyun-Bum Kim

Towards Automatic Identification of Self-Reported COVID-19 Tweets: Introducing a Multilingual Manually Annotated Dataset, Baseline Systems and Exploratory Evaluations 3736
Ramya Tekumalla, Luis Alberto Robles Hernandez, Juan M. Banda

Energy-Aware Query Processing: A Case Study on Join Reordering 3743
Ladjel Bellatreche, Fouad Djellali, Wojciech Macyna, Carlos Ordonez

A Methodology and a Toolbox to Explore Dataset Related to the Environmental Impact of HTTP Requests 3753
Christophe Cérin, Mathilde Jay, Laurent Lefèvre, Denis Trystram

Explainable Deep Learning for Smart Grid Stability Detection 3763
Giovanni Ciaramella, Fabio Martinelli, Francesco Mercaldo, Antonella Santone

Natural Language Processing to Understand Human Activities Impacted by Hydroelectric Energy Projects 3770
Keshava Pallavi Gone, Yan Chen, Michael Smit

Wireless Sensor Networks Calibration Using Attention-Based Gated Recurrent Units for Air Pollution Monitoring 3779
Shagufta Henna, Asif Yar, Kazeem Saheed, Paulson Grigarichan

The Forecastability of Underlying Building Electricity Demand from Time Series Data 3785
Mohamad Khalil, A. Stephen McGough, Hussain Kazmi, Sara Walker

Explainable Long-Term Forecasting of Air Pollution 3794
Abdul Mutakabbir, Samuel A. Ajila

Counterfactual Analysis: A Case Study on Impact of External Events on Building Energy Consumption 3803
Carolina Minami Oguchi, Dipak Ghosal, Alex Sim, Kesheng Wu

Forecasting Resource Demand for Dynamic Datacenter Sizing in Telco Infrastructures 3813
Dimitra Paranou, Angelos Pentelas, Dimitris Katsiros, Konstantinos Maidatsis, George Giannopoulos, Evangelos Angelou, Nikos Anastopoulos, George Papastefanatos

Is Active Learning Green? an Empirical Study 3823
Shirin Salehi, Anke Schmeink

Detection of Methane Plumes Using Sentinel-2 Satellite Images and Deep Neural Networks Trained on Synthetically Created Label Data 3830
Maciel Zortea, João Lucas De Sousa Almeida, Levente Klein, Alberto Costa Nogueira Junior

Predictive Modeling of Run Time for Model and Data Distributed Inferencing Using Gradient Boosting Regression 3840
Scott Brown, David Harman, Cleon Anderson, Matthew Dwyer

Thoughts on Edge Sensing and Processing 3847
J. Michael Barton

Measuring Data Transmissions from the Edge for Distributed Inferencing with gRPC 3853
Scott Brown, David Harman, Cleon Anderson, Matthew Dwyer

Deep Learning Acceleration at the Resource-Constrained Tactical Edge 3857
Billy Geerhart, Venkat R. Dasari, Peng Wang, Brian Rapp

Anomaly Detection as a Data Reduction Approach for Test Event Analysis at the Edge..... 3863
Mariya Occorso, Michael An, Robert Olsen, Vincent Perry

Real-Time Camera-To-Lidar Calibration for Autonomous Robotic Systems at the Edge..... 3868
Ju Wang, Venkat R. Dasari, Billy Geerhart, Brian Rapp, Peng Wang

Performance Evaluation of 3D Point Cloud Matching with Short-Range Dense 2D Rang..... 3874
Ju Wang, Ahmed F. Mohammed, Grant Patterson

Enhancing Edge Intelligence with Highly Discriminant LNT Features 3880
Xinyu Wang, Vinod K. Mishra, C.-C. Jay Kuo

Navigating the Future Data-Driven Automation Tools: State-Of-The-Art and Research Roadmap for Digital Twins of Energy Systems..... 3888
Ali Aghazadeh Ardebili, Antonella Longo, Antonio Ficarella

A Big Data Platform for the Management of Local Energy Communities Data 3898
Sonia Bergamaschi, Luca Gagliardelli

Secure and Energy Efficient Filtered Over-The-Air Internet of Things Setup in a Wireless Mesh Network for Firmware Freshness 3904
Lorenzo Carnevale, Annamaria Ficara, Gianluca Catalfamo, Antonino Galletta, Maria Fazio, Massimo Villari

Enhanced Segmentation of PV Arrays in Infrared Images Using an Improved SegFormer Approach 3914
Wanghu Chen, Shengfang Jin, Yihua Luo, Jing Li

An Integrated ML-Ops Framework for Automating AI-Based Photovoltaic Forecasting 3921
Symeon Chorozoglou, Elissaios Sarmas, Vangelis Marinakis

Embedding Climate Dynamics and Prediction with Deep Learning for Wind Power Forecasting: Short-Term to Long-Term Perspective 3929
Hana Kim, Yong Oh Lee, Changsoo Ok, Dongkyun Kim, Seungyup Baek

A Power Quality Event Classifier Based on Convolutional Neural Network..... 3936
Francesco Lamonaca, Domenico Luca Carnì

A Deep Learning Approach for Climate Parameter Estimations and Renewable Energy Sources..... 3942
Camilla Lops, Mariano Pierantozzi, Luciano Caroprese, Sergio Montelpare

Physics Fundamentals of an Internet of Nanomedicine Things..... 3952
Huber Nieto-Chaupis

Kirchhoff's Voltage Law at Neural Synapse from Electric Currents Based at Neurotransmiters..... 3956
Huber Nieto-Chaupis

A Focused Review of ANN-Based Models for Predicting Absorption Maxima (λmax) of Dyes..... 3961
Geeta Rani, Neeraj Tomar, Vijaypal Singh Dhaka, Praveen K. Surolia, Eugenio Vocaturo, Ester Zumpano

Windfarm Forced Oscillation Detection Using Hyperdimensional Computing 3965
Shyam Yathirajam, Arash Peighambari, Ruben Roberts, Hamed Nademi, Sreedevi Gutta, Justin Morris, Ali Ahmadinia

Supply-Demand Analysis of Urban Amenities Based on Walking Accessibility 3973
Kaloyan Karamitov, Dessislava Petrova-Antonova, Emil Hristov, Milena Borukova

CkanFAIR: A Digital Tool for Assessing the FAIR Principles 3980
Matteo Lia, Davide Damiano Colella

Identifying Key Factors in Designing Data Spaces for Urban Digital Twin Platforms: A Data Driven Approach 3985
Cristian Martella, Angelo Martella, Amro Issam Hamed Attia Ramadan

Cases Vs Deaths: Which Indicators to Assess the Effectiveness of Non-Pharmaceutical Interventions During Covid-19 Pandemic? 3995
Divya Pragna Mulla, Mario Bochicchio, Antonella Longo

Advancing Resilience in Green Energy Systems: Comprehensive Review of AI-Based Data-Driven Solutions for Security and Safety 4002
Amro Issam Hamed Attia Ramadan, Ali Aghazadeh Ardebili, Antonella Longo, Antonio Ficarella

Digital Twin Supporting Blockchain in Construction Supply Chain: From Bio-Based Materials to Energy Efficiency and L.C.A.. 4011
Angelica Rocco

Digital Twin Space: The Integration of Digital Twins and Data Spaces..... 4017
Alessandra Somma, Alessandra De Benedictis, Marco Zappatore, Cristian Martella, Angelo Martella, Antonella Longo

Generating Large-Scale Human Flow Datasets from Measured Pedestrian Movement Data and Simulation 4026
Mei Takeda, Masaki Onishi

Vision Paper: Uncovering Illegal Firearm Transactions in Cryptocurrency Networks 4031
Anastasia Kassiani Blitsi, Georgios Stavropoulos, Konstantinos Votis

Visual Data Engineering for Conflict and Terrorism Prediction..... 4036
Antonella Calò, Matteo Lia, Marco Zappatore, Antonella Longo

Harvesting Large Textual and Multimedia Data to Detect Illegal Activities on Dark Web Marketplaces 4046
Georgia Chatzimarkaki, Sophia Karagiorgou, Mariza Konidi, Dimitrios Alexandrou, Thanassis Bouras, Spyridon Evangelatos

The Nexus Between Big Data Analytics and the Proliferation of Fake News as a Precursor to Online and Offline Criminal Activities..... 4056
Spyridon Evangelatos, Thanasis Papadakis, Nikolaos Gousetis, Christos Nikolopoulos, Petrina Troulitaki, Nikos Dimakopoulos, George Bravos, Michael V. Lo Giudice, Ali Shadma Yazdi, Alberto Aziani

General Risk Index : A Measure for Predicting Violent Behavior Through Written Communication..... 4065
Lisa Kaati, Amendra Shrestha, Nazar Akrami

Self-Supervised Visual Learning for Analyzing Firearms Trafficking Activities on the Web 4071
Sotirios Konstantakos, Despina Ioanna Chalkiadaki, Ioannis Mademlis, Adamantia Anna Rebolledo Chrysochoou, Georgios Th. Papadopoulos

Visual Inspection for Illicit Items in X-Ray Images Using Deep Learning..... 4081
Ioannis Mademlis, Georgios Batsis, Adamantia Anna Rebolledo Chrysochoou, Georgios Th. Papadopoulos

Unveiling Conversational Patterns: Intent Classification and Clustering in Reddit's Firearm Trade Community 4090
Maria Makrynioti, Chaido Porlou, Anastasios Alexiadis, Georgios Stavropoulos, Konstantinos Votis, Dimitrios Tzovaras

Investigating Visual Analytics Against Terrorist Financing in Dark Web Marketplaces 4100
Kyriaki Miniadou, Eirini Kyriakou, Spyridon Tzagkarakis, Asterios Leonidis, Georgios Th. Papadopoulos, Constantine Stephanidis

Cybersecurity in Public Space: Leveraging CNN and LSTM for Proactive Multivariate Time Series Classification 4110
Aimen Ahmed Al Odaini, Francesco Zola, Lander Segurola-Gil, Amaia Gil-Lertxundi, Carmen D'Andrea

GINN: Towards Gender InclusioNeural Network 4119
Matteo Berta, Bartolomeo Vacchetti, Tania Cerquitelli

Mitigating Representation Bias in Data Transformations: A Constraint-Based Optimization Approach 4127
Barbara Catania, Giovanna Guerrini, Ziad Janpih

Exploring Factors Influencing the Adoption of In-Home Respite Services: A Data Science Approach 4137
Stephen CC Cheng, Clio Yuen Man Cheng, Dah Ming Chiu, Alice Ming Lin Chong, Vivian W Q Lou

Detecting Anomalous Suppliers in Electronic Procurement Systems Using GAAN 4147
Eduardo Cruz, Natalia Ramirez, Emanuel Parra, Daniel Ochoa

A Perspective on Data Categorization with Regard to Equity, Diversity, and Inclusion in Data Science 4156
Cécile Favre, Marie Vialaret

Intentional Youth Development Activities and Peer Effects in a Gang Prevention Program 4166
Haoran Jia, Haoyu Li, Xinyue Li, Xiaoxian Shen, Yichen Wang, Zichun Liao, Andrea L. Bertozzi, P. Jeffrey Brantingham, Jona Lelmi

Developing a Large-Scale Language Model to Unveil and Alleviate Gender and Age Biases in Australian Job Ads 4176
Ruochen Mao, Liming Tan, Rezza Moieni

Assessing Relative Contribution of Environmental, Behavioural and Social Factors on Life Satisfaction Via Mobile App Data 4186
Gregor Milligan, John Harvey, Liz Dowthwaite, Elvira Perez Vallejos, Georgiana Nica-Avram, James Goulding

Recommendation Fairness and Where to Find It: An Empirical Study on Fairness of User Recommender Systems 4195
Antonela Tommasel, Ira Assent

Aiming to Minimize Alcohol-Impaired Road Fatalities: Utilizing Fairness-Aware and Domain Knowledge-Infused Artificial Intelligence 4205
Tejas Venkateswaran, Sheikh Rabiul Islam, Md Golam Moula Mehedi Hasan, Mohiuddin Ahmed

Introducing Computational Thinking in Middle-Schools Using a Culturally-Responsive Game Through a Researcher-Practitioner Partnership 4213
Ismael Villanueva-Miranda, Katherine Mortimer, Monika Akbar, Romelia Rodriguez Reyes, Cynthia Ontiveros, Scott Gray, Pedro Delgado, Victor Medrano, Melissa Anderson, Jacob Ramirez, Jesus Vazquez

PSNR-Aware Quantization for DCT-Based Lossy Compression 4223
Jiaxi Chen, Seung Woo Son

Exploring Wavelet Transform Usages for Error-Bounded Scientific Data Compression 4233
Jiajun Huang, Jinyang Liu, Sheng Di, Yujia Zhai, Zizhe Jian, Shixun Wu, Kai Zhao, Zizhong Chen, Yanfei Guo, Franck Cappello

GLEm-Net: Unified Framework for Data Reduction with Categorical and Numerical Features 4240
Francesco De Santis, Danilo Giordano, Marco Mellia, Alessia Damilano

GAME: Improving Efficiency and Effectiveness of Knowledge-Graph Rule Mining Via Data Reduction 4248
Kara Schatz, Alexander Tropsha, Rada Chirkova

Mutually Exclusive Learning for Generators with Multi-Label Classifiers 4258
Digya Acharya, Hera Siddiqui, Eduardo Pasiliao, Chaity Banerjee

Novel NBA Fantasy League Driven by Engineered Team Chemistry and Scaled Position Statistics 4268
Ganesh Arkanath, Nishad Gupta, Hasan Kurban, Parichit Sharma, K R Madhavan, Elham Khorasani Buxton, Mehmet M. Dalkilic

Tensor Space Model-Based Textual Data Augmentation for Text Classification 4276
Minsuk Chang, Han-Joon Kim

ASI: Accuracy-Stability Index for Evaluating Deep Learning Models 4284
Wei Dai, Daniel Berleant

Enabling Cross-Language Data Integration and Scalable Analytics in Decentralized Finance 4290
Conor Flynn, Kristin P. Bennett, John S. Erickson, Aaron Green, Oshani Seneviratne

MIAE: A Mobile Application Recommendation Method Based on a NTK Model 4300
Jiahui Han, Qufei Zhang, Xiaoying Yang, Jinyi Wang

FedSym: Unleashing the Power of Entropy for Benchmarking the Algorithms for Federated Learning 4308
S. Ensiye Kiyamousavi, Boris Kraychev, Ivan Koychev

LLM-TAKE: Theme-Aware Keyword Extraction Using Large Language Models 4318
Reza Yousefi Maragheh, Chenhao Fang, Charan Chand Irugu, Parth Parikh, Jason Cho, Jianpeng Xu, Saranyan Sukumar, Malay Patel, Evren Korpeoglu, Sushant Kumar, Kannan Achan

GPT in Data Science: A Practical Exploration of Model Selection 4325
Nathalia Nascimento, Cristina Tavares, Paulo Alencar, Donald Cowan

Combining Block Bootstrap with Exponential Smoothing for Reinforcing Non-Emergency Urban Service Prediction 4335
Kshira Sagar Sahoo, Shivam Krishana, Monowar Bhuyan

Tissue-Specific Color Encoding and GAN Synthesis for Enhanced Medical Image Generation..... 4344
Yu Shi, Hannah Tang, Jianxin Sun, Xinyan Xie, Huijing Du, Dandan Zheng, Chi Zhang, Hongfeng Yu

Effect of Varied Datasets on Training of a Segmentation Model Used in Visual Navigation..... 4350
Marin Wada, Miho Adachi, Ryusuke Miyamoto

Does a Dense Point Cloud for Training Data Generation Improve Segmentation Accuracy?..... 4356
Marin Wada, Hiroaki Sudo, Miho Adachi, Ryusuke Miyamoto

Expressway: Prioritizing Edges for Distributed Evaluation of Graph Queries..... 4362
Abbas Mazloumi, Mahbod Afarin, Rajiv Gupta

Adapting Knowledge Graphs to Edge Computing Devices 4372
Joffrey De Oliveira, Christophe Callé, Olivier Curé

On the Pursuit of Graph Embedding Strategies for Individual Mobility Networks..... 4382
Omid Isfahani Alamdari, Mirco Nanni, Agnese Bonavita

Multi-Modal Medical Information Based Data Mining for Expression and Characteristic Pattern Prediction of TP53 in Endometrial Carcinoma..... 4392
Tianming Du, Chen Li, Tao Jiang, Jinzhu Yang, Marcin Grzegorzek, Hongzan Sun

Predicting PD-L1 Status of Esophageal Cancer from H&E Images Based on FusedNet Model 4397
Minghe Gao, Chen Li, Jingyu Zhang, Haoyuan Chen, Weiming Hu, Hechen Yang, Liyu Shi, Yujie Jing, Shuaiyi Tian, Hongzan Sun, Marcin Grzegorzek, Xiaoyan Li

Continuous Prompt for Chemical Language Model Aided Anticancer Synergistic Drug Combination Prediction..... 4406
Guannan Geng, Lingling Zhao, Chunyu Wang, Xiaoyan Liu, Junjie Wang

A New Prediction Algorithm Based on Local Neighbor Closeness and Influence of Node 4413
Haiping Ding, Liang Wei, Fazu Li, Feng Hu, Pengyue Li, Junjie Chen

Dermoscopic Image Classification Using Attention Mechanism and Ensemble Learning Approaches..... 4424
Shanchuan Huang, Hongwei Lei, Liuhan Jin, Jinzhu Yang, Tao Jiang, Yudong Yao, Marcin Grzegorzek, Chen Li

Multi-Scale Heterogeneous Graph Contrastive Learning..... 4432
Mingyuan Li, Lei Meng, Zhonglin Ye, Haixing Zhao, Yuzhi Xiao, Shujuan Cao

Clustering and GNN Prediction with DrugMatrix..... 4442
Jiaji Ma, Guojing Cong, Scott Auerbach

Proximity Network for Visualizing Infection Risks of Pedestrian Behavior at Large-Scale Events 4450
Sayaka Morikoshi, Ryo Niwa, Shunki Takami, Masaki Onishi, Takayuki Itoh

Graph Coordinates and Conventional Neural Networks - An Alternative for Graph Neural Networks 4456
Zheyi Qin, Randy Paffenroth, Anura P. Jayasumana

ECA-RetinaNet: A Novel Self-Attention RetinaNet for Environmental Microorganism Image Object Detection..... 4466
Hechen Yang, Jinzhu Yang, Tao Jiang, Xin Zhao, Jinghua Zhang, Peng Zhao, Ao Chen, Qianqing Nie, Marcin Grzegorzek, Chen Li

The Application of Long Short-Term Memory and Transformers for Music Generation..... 4475
B Shanmukh Krishna, Satya Sai Roopa Sree Chiluvuri, Sai Sri Harshini Kosuri, Sai Sree Harsha Chimbili

Music Generation Using Human-In-The-Loop Reinforcement Learning 4479
Aju Ani Justus

Automatic Time Signature Determination for New Scores Using Lyrics for Latent Rhythmic Structure 4485
Callie C. Liao, Duoduo Liao, Jesse Guessford

AI-Algorithmically-Generated Song with Lyrics 4495
Callie Liao

Musical Elements Enhancement and Image Content Preservation Network for Image to Music Generation 4497
Wenzhao Liu, Dechao Meng

Things, Objects, Subjects and Stuff: IoMuSt and Ubimus Perspectives on AI 4503
Marcello Messina, Luzilei Aliel

Snake in the Labyrinth: Scenes from the Machine's Deep Q-Learning Experience.....4511
Jeffrey M. Morris

Equipping Pretrained Unconditional Music Transformers with Instrument and Genre Controls..... 4512
Weihan Xu, Julian McAuley, Shlomo Dubnov, Hao-Wen Dong

Let the 'Perfect Voice' Always Be There, Positive Affirmation and Negative Anxiety from Online Commentary 4518
Yingzixuan Zhang, Kaixuan Niu, Zhian Zhao

Assessing the Environmental Impact of Plant-Based Diets: A Comprehensive Analysis..... 4521
Blagica Golubova, Fjola Fetaji, Jovana Dobreva, Milena Trajanoska, Ana Todorovska, Riste Stojanov, Dimitar Trajanov

Who Consumes Anthocyanins and Anthocyanidins? Mining National Retail Data to Reveal the Influence of Socioeconomic Deprivation and Seasonality on Polyphenol Dietary Intake 4530
John Harvey, Gavin Long, Roberto Mansilla, Simon Welham, Peter Rose, Michelle Thomas, Gregor Milligan, Elizabeth Dolan, Joanne Parkes, Kuzivakwashe Makokoro, James Goulding

Methodology for Food Prices Forecasting 4539
Dimitar Peshevski, Ana Todorovska, Filip Trajkovikj, Nikola Hristov, Milena Trajanoska, Jovana Dobreva, Riste Stojanov, Dimitar Trajanov

Empowering Cover Crop Decision Support with Visualization and Provenance Enhancement 4548
Sujan Shrestha, Jianxin Sun, Katja Koehler-Cole, Andrea Basche, Hongfeng Yu

A Comprehensive Study of Food Prices and Food Fraud in the European Union..... 4557
Filip Trajkovikj, Ana Todorovska, Dimitar Peshevski, Lina Nakova, Milena Trajanoska, Jovana Dobreva, Riste Stojanov, Dimitar Trajanov

Neural Networks Voting for Projection Based Ensemble Classifiers 4567
Panagiotis Anagnostou, Petros Barmpas, Sotiris K. Tasoulis, Spiros V. Georgakopoulos, Vassilis P. Plagianakos

Mask-CscGAN for Realistic Synthetic Cell Generation 4575
Panagiotis Antoniadis, Christina Sartzetaki, Nick Antonopoulos, Pantelis Papageorgiou, Aigli Korfiati, Vassilis Pitsikalis

A Graph-Based Approach to Integrate Large-Scale Drug and Protein Data for Alzheimer's Disease Drug Repurposing 4584
Georgios N. Dimitrakopoulos, Konstantinos Lazaros, Marios Krokidis, Themis Exarchos, Aristidis G. Vrahatis, Panagiotis Vlamos

Efficient Time Series Classification Based on Learning Similar Trend Features in the Same Class Sequences 4588
Guohui Ding, Zhaoyi Yuan, Wenjing Tang, Chao Jiang

Speeding Up Unsupervised Learning Through Early Stopping 4597
Andris Docaj, Yu Zhuang

Enhancing Disease Diagnosis: A CNN-Based Approach for Automated White Blood Cell Classification 4606
Athanasios Kanavos, Orestis Papadimitriou, Alexios Kaponis, Manolis Maragoudakis

Advanced Big Data Analysis for Deciphering the Role of Protein Misfolding and Interactions in the Pathogenesis of Alzheimer's Disease 4614
Marios G. Krokidis, Georgios N. Dimitrakopoulos, Themis P. Exarchos, Aristidis G. Vrahatis, Panagiotis Vlamos

Decoding Gender on Social Networks: An In-Depth Analysis of Language in Online Discussions Using Natural Language Processing and Machine Learning 4618
Gerasimos Vonitsanos, Andreas Kanavos, Phivos Mylonas

Towards Automated Regulatory Compliance Verification in Financial Auditing with Large Language Models 4626
Armin Berger, Lars Hillebrand, David Leonhard, Tobias Deußer, Thiago Bell Felix De Oliveira, Tim Dilmaghani, Mohamed Khaled, Bernd Kliem, Rudiger Loitz, Christian Bauckhage, Rafet Sifa

Model-As-A-Service (MaaS): A Survey 4636
Wensheng Gan, Shicheng Wan, Philip S. Yu

Unsupervised Style Transfer of Modern Hebrew Using Generative Language Modeling and Zero-Shot Prompting 4646
Pavel Kaganovich, Ophir Münz-Manor, Elishai Ezra Tsur

Wealth of Nations, Wealth of Data: How GDP Shapes Diverse Large Language Models Like ChatGPT : Interviewing Assorted Open Source Generative AI Models 4654
Alex Kaplunovich

Comparing Generative Chatbots Based on Process Requirements: A Case Study 4664
Luis Fernando Lins, Nathalia Nascimento, Paulo Alencar, Toacy Oliveira, Donald Cowan

GPT-In-The-Loop: Supporting Adaptation in Multiagent Systems 4674
Nathalia Nascimento, Paulo Alencar, Donald Cowan

Generating Prototypes for Contradiction Detection Using Large Language Models and Linguistic Rules 4684
Maren Pielka, Svetlana Schmidt, Rafet Sifa

Hallucination-Minimized Data-To-Answer Framework for Financial Decision-Makers 4693
Sohini Roychowdhury, Andres Alvarez, Brian Moore, Marko Krema, Maria Paz Gelpi, Punit Agrawal, Federico Martín Rodríguez, Ángel Rodríguez, José Ramón Cabrejas, Pablo Martínez Serrano, Arijit Mukherjee

TRANSQLATION: TRANsformer-Based SQL RecommendATION 4703
Shirin Tahmasebi, Amir H. Payberah, Ahmet Soylu, Dumitru Roman, Mihhail Matskin

Enabling Dataspaces Using Foundation Models: Technical, Legal and Ethical Considerations and Future Trends..... 4712
Mohan Timilsina, Samuele Buosi, Ping Song, Yang Yang, Rafiqul Haque, Edward Curry

Anaphoric Ambiguity Resolution in Software Requirement Texts..... 4722
Sanaz Mohammad Jafari, Savas Yildirim, Mucahit Cevik, Ayse Basar

Toward Open and Sustainable Data Platform for Accessible Pedestrian Network..... 4731
Masahiro Bessho, Tomomori Usaka, Ken Sakamura

ITDT: International Testbed for Dataspace Technology..... 4740
Itsuki Matsunaga, Takashi Michikata, Noboru Koshizuka

Applying Homomoprhic Encryption to Data Spaces Takashi Michikata. 4748
Takashi Michikata, Yusuke Sasaki, Noboru Koshizuka

Hierarchical Rule Compliance Check Method for Distributed Query..... 4754
Divya Singh, Mika Takata, Toshihiko Kashiyama, Satoshi Katsunuma, Norifumi Nishikawa

Unveiling Engagement in Virtual Classrooms: A Multimodal Analysis..... 4761
Diletta Chiaro, Daniela Annuziata, Stefano Izzo, Francesco Piccialli

Early ChatGPT User Portrait Through the Lens of Data 4770
Yuyang Deng, Ni Zhao, Xin Huang

Large Language Models in Education: Vision and Opportunities 4776
Wensheng Gan, Zhenlian Qi, Jiayang Wu, Jerry Chun-Wei Lin

Robust Category Recognition Based on Deep Templates for Educational Mobile Applications 4786
Marzena Halama, Katarzyna Filus, Joanna Domanska

A Hybrid Corpus Based Fine-Grained Semantic Alignment Method for Pre-Trained Language Model of Ancient Chinese Poetry..... 4794
Jie Hong, Tingting He, Jie Mei, Ming Dong, Zheming Zhang, Xinhui Tu

HierAct: A Hierarchical Model for Human Activity Recognition in Game-Like Educational Applications..... 4802
Nur Kelesoglu, Katarzyna Filus, Joanna Domanska

Game Based Virtual Reality Platform for Geometry..... 4812
Himangshu Sarma, Mahaboob Shaik, Surya Teja Tangirala, Satyam Kumar Singh

A Reflection on Cybersecurity Indigenous Educational Experiences 4819
Kayleigh Tanner, Sarah Plosker, Gautam Srivastava

An Empirical Study of Covid-19 Effect on Health Care Workers' Career Development..... 4826
Kutay Acar, Günce Keziban Orman, Sultan Nezihe Turhan

Incorporating Emotions into Health Mention Classification Task on Social Media 4834
Olanrewaju Tahir Aduragba, Jialin Yu, Alexandra I. Cristea

Explainable NLP Model for Predicting Patient Admissions at Emergency Department Using Triage Notes 4843
Emilien Arnaud, Mahmoud Elbattah, Pedro A. Moreno-Sánchez, Gilles Dequen, Daniel Aiham Ghazali

On the Usage of Benford's Law on Coronavirus Data Statistics 4848
Sena Atakan, Günce Keziban Orman

Toward Training NLP Models to Take into Account Privacy Leakages 4854
Gaspard Berthelier, Antoine Boutet, Antoine Richard

Impact and Analysis of Long COVID on Cardiovascular Diseases 4863
Daniel Chia, Priya Deshpande

Visual Analytics and Exploration of Calcium Transient Imaging Data Using Event-Based Clustering 4868
Shiaofen Fang, Michal Lange, Haoying Fu, Yingying Chen, Yaoying Ma

Learning Cataract Severity Using the Contrast Sensetivity Scale: A Thick Data Approach 4875
Jinan Fiaidhi, Sabah Mohammed

Multi-Task Learning Using Facial Features for Mental Health Screening 4881
Ricardo Flores, Avantika Shrestha, Ml Tlachac, Elke A. Rundensteiner

A Classification Approach to Detect Social Support in Online Health Communities 4891
Hai Nguyen, Thanaa Ghanem

Deep Learning Based Driving Posture Stability Analysis for People with Mobility Challenges 4900
Ruoqi Huang, Mai Narasaki-Jara, Tingting Chen

Classification of Decompression Sickness Signs and Symptoms Using Cluster Analysis 4907
Jean Pierre Imbert, Salih Murat Egi, Tamer Özyigit

Autocompletion of Chief Complaints in the Electronic Health Records Using Large Language Models 4912
K M Sajjadul Islam, Ayesha Siddika Nipu, Praveen Madiraju, Priya Deshpande

Combat COVID-19 at National Level Using Risk Stratification with Appropriate Intervention 4922
Xuan Jin, Kar Way Tan

Extending Healthy Life Expectancy: Frailty Detection Model Using Smart Meter and Machine Learning 4931
Kijun Kim, Noboru Koshizuka

Integrating a PICO Clinical Questioning to the QL4POMR Framework for Building Evidence-Based Clinical Case Reports 4940
Sabah Mohammed, Jinan Fiaidhi, Rahul Kudadiya

Precise Temperature Control for Sensitive Medicines 4948
Mehmet Kerim Peker, Ahmet Teoman Naskali, Sanem Kiliçaslan, Nuri Özbey

Predicting Mental Health Disorders Post Long COVID Diagnosis Using Advanced Machine Learning Techniques 4954
Manoj Purohit, Praveen Madiraju

Is One Label All You Need? Single Positive Multi-Label Training in Medical Image Analysis 4963
Helen Schneider, Priya Priya, David Biesner, Rebecca Kador, Yannik C. Layer, Maike Theis, Sebastian Nowak, Alois M. Sprinkart, Ulrike I. Attenberger, Rafet Sifa

Real-Time Heart Rate Monitoring Via Wi-Fi Signal 4973
Hasan Ali Solgun, Alp Ozcan, Ozgun Pinarer

What is Hiding in Medicine's Dark Matter? Learning with Missing Data in Medical Practices..... 4979
Neslihan Suzen, Evgeny M. Mirkes, Damian Roland, Jeremy Levesley, Alexander N. Gorban, Tim J. Coats

Comparison of Machine Learning Classifiers on Integrated Transcriptomic Data..... 4987
Irene Testa, Giuseppe Prencipe, Corrado Priami, Alina Sîrbu

Knowledge Graphs, Clinical Trials, Dataspace, and AI: Uniting for Progressive Healthcare Innovation..... 4997
Mohan Timilsina, Saeed Alsamhi, Rafiqul Haque, Conor Judge, Edward Curry

Leveraging Graph Databases for Enhanced Healthcare Data Management: A Performance Comparison Study 5007
Sultan N. Turhan

More than Just a Diagnosis: A Multi-Task Approach to Analyzing Bipolar Disorder on Reddit Via DeMHeM 5014
Rocco Zhang, Shailik Sarkar, Abdulaziz Alhamadani, Chang-Tien Lu

Effective Prediction of Energy Consumption in Automated Guided Vehicles with Recurrent and Convolutional Neural Networks 5024
Pawel Benecki, Daniel Kostrzewa, Piotr Grzesik, Bohdan Shubyn, Jia-Hao Syu, Jerry Chun-Wei Lin, Vaidy Sunderam, Dariusz Mrozek

The Calibration of Single Beam Distance Sensors Based on Machine Learning Methods 5031
Piotr Biernacki, Adam Ziebinski, Jia-Hao Syu, Jerry Chun-Wei Lin

NetSat: Network Saturation Adversarial Attack 5038
Katarzyna Filus, Joanna Domanska

Reliability Analysis of M2M Cyclic Data Transfer Based on Ad-Hoc Wireless P2P Link 5048
Piotr Gaj, Ireneusz Smolka, Jacek Stój

Estimating the AGV Load and a Battery Lifetime Based on the Current Measurement and Random Forest Application 5057
Krzysztof Hanzel, Damian Grzechca, Adam Ziebinski, Lukas Chruszczyk, Artur Janus

A Fuzzy Optimized Route Selection Framework for Autonomous Vehicles Using V-NDN 5064
Anu Kaushik, Rasmeet S Bali, Gautam Srivastava

General Concepts in Swarm of Drones Control: Analysis and Implementation..... 5070
Dariusz Marek, Marcin Paszkuta, Jakub Szygula, Piotr Biernacki, Adam Domanski, Marta Szczygiel, Marcel Król, Konrad Wojciechowski

Testing Quality of Service of Communication System for AGV Fleet with Software-Defined Network..... 5078
Karol Marszalek, Adam Domanski, Rafal Cupek, Marek Drewniak

Predicting AGV Battery Cell Voltage Using a Neural Network Approach with Preliminary Data Analysis and Processing 5087
Olena Pavliuk, Mykola Medykovskyy, Tomasz Steclik

Lightweight CNN Based on Spatial Features for a Vehicular Damage Detection System 5097
Dawid Polap, Antoni Jaszcz, Katarzyna Prokop, Gautam Srivastava

Unsupervised Learning for Depth Estimation in Unstructured Environments 5102
Pian Qi, Fabio Giampaolo, Edoardo Prezioso, Francesco Piccialli

Analysis of Time Management Models in AGV Wireless Communication 5110
Kirill Rukkas, Piotr Gaj

Visual Examination of Relations Between Known Classes for Deep Neural Network Classifiers 5118
Lukasz Sobczak, Katarzyna Filus, Marzena Halama, Joanna Domanska

Convergent Wireless Real-Time Communicaiton Using EtherCAT and OPC UA for AGV Application with Big Data Sets Acquisition 5127
Jacek Stój, Piotr Gaj

UniMHe: Unified Multi Hyperedge Prediction a Case Study on Crime Dataset..... 5134
Melike Yildiz Aktas, Lulwah Alkulaib, Chang-Tien Lu

Balancing the Scales: HyperSMOTE for Enhanced Hypergraph Classification 5140
Lulwah Alkulaib, Chang-Tien Lu

IDMU: Impact Driven Machine Unlearning 5146
Shubhi Asthana, Bing Zhang, Ruchi Mahindru, Indervir Singh Banipal, Pawan Chowdhary

Choosing the Right Time to Learn Evolving Data Streams..... 5156
Alessio Bernardo, Emanuele Della Valle, Albert Bifet

Using Human Mobility Patterns to Forecast Outliers in Citizen Complaints Data 5166
Vincent Bolta, Marwan Hassani

Artificial Intelligence for Spectral Analysis: Challenges and Opportunities 5176
Anqi Chu, Xiang Xie, Wilhelm Stork

Fine-Tuned Generative LLM Oversampling Can Improve Performance Over Traditional Techniques on Multiclass Imbalanced Text Classification 5181
Nicolas Antonio Cloutier, Nathalie Japkowicz

Effective and Efficient Big OLAP Data Cube Compression in Mobile Cloud Environments: The IQTS Algorithm..... 5187
Alfredo Cuzzocrea, Mojtaba Hajian

F-TBDA: A Frequency-Based Temporal Big Data Analytics Technique for Mining and Analyzing Quality-Of-Life Indicators of Cancer Patients..... 5197
Alfredo Cuzzocrea, Geertruida H. De Bock, Willemijn J. Maas, Selim Soufargi

Machine-Learning-Based Multidimensional Big Data Analytics Over Clouds Via Multi-Columnar Big OLAP Data Cube Compression 5206
Alfredo Cuzzocrea, Abderraouf Hafsaoui, Carson K. Leung

DeepScreen: Boosting Depression Screening Performance with an Auxiliary Task 5213
Ricardo Flores, Avantika Shrestha, Elke A. Rundensteiner

Stabilizing Adversarial Training for Generative Networks 5223
Walter Gerych, Kevin Hickey, Thomas Hartvigsen, Luke Buquicchio, Abdulaziz Alajaji, Kavin Chandrasekaran, Hamid Mansoor, Emmanuel Agu, Elke Rundensteiner

Cross-Silo Federated Learning Across Divergent Domains with Iterative Parameter Alignment 5233
Matt Gorbett, Hossein Shirazi, Indrakshi Ray

Neurosymbolic Knowledge Distillation 5243
Himel Das Gupta, Victor S. Sheng

UCE-FID: Using Large Unlabeled, Medium Crowdsourced-Labeled, and Small Expert-Labeled Tweets for Foodborne Illness Detection 5250
Ruofan Hu, Dongyu Zhang, Dandan Tao, Huayi Zhang, Hao Feng, Elke Rundensteiner

Diverse Data Expansion with Semi-Supervised k-Determinantal Point Processes 5260
Simon Johansson, Ola Engkvist, Morteza Haghir Chehreghani, Alexander Schliep

NASEREX: Optimizing Early Exits Via AutoML for Scalable Efficient Inference in Big Image Streams 5266
Aakash Kapoor, Rajath Elias Soans, Soham Dixit, Pradeep Ns, Brijraj Singh, Mayukh Das

Spider4SPARQL: A Complex Benchmark for Evaluating Knowledge Graph Question Answering Systems..... 5272
Catherine Kosten, Philippe Cudré-Mauroux, Kurt Stockinger

Multi-Item Continuous Influence Maximization..... 5282
Jianjun Luo, Ryan A. Rossi, Xiangnan Kong, Yanhua Li

Student Success Analysis for Minority Students in Higher Education 5292
Shamiha Binta Manir, Priya Deshpande

Modular Deep Learning for Big Data: Motivation, Challenges, and Approaches..... 5299
Samiyuru Menik, Lakshmish Ramaswamy

DataCockpit: A Toolkit for Data Lake Navigation and Monitoring Utilizing Quality and Usage Information..... 5305
Arpit Narechania, Surya Chakraborty, Shivam Agarwal, Atanu R Sinha, Ryan A. Rossi, Fan Du, Jane Hoffswell, Shunan Guo, Eunyee Koh, Alex Endert, Shamkant Navathe

Hierarchical Uncertainty Aggregation and Emphasis Loss for Active Learning in Object Detection.....5311
Tai Nguyen, Khoa Nguyen, Thanh Nguyen, Tri Nguyen, Anh Nguyen, Karrman Kim

Improved Tactical Decision Making and Control Architecture for Autonomous Truck in SUMO Using Reinforcement Learning 5321
Deepthi Pathare, Leo Laine, Morteza Haghir Chehreghani

Speech-Based Detection of In-Car Escalating Arguments to Prevent Distracted Driving..... 5330
Francesco Pistolesi, Michele Baldassini, Beatrice Lazzerini

Uncertainty-Aware Optimisation for Sustainable Multimedia Event Processing in Big Data Streams..... 5338
Felipe Arruda Pontes, Michael Schukat, Edward Curry

Online Hierarchical Multi–label Classification 5346
Wenting Qi, Charalampos Chelmis

Scaling Personalized Machine Learning Through DTW Clustering: Predicting Glycemia Levels as an Example 5356
Abdelmounaam Rezgui

Towards Big Data Analytics Over Mobile User Data Using Machine Learning 5365
Sabrina Ichou, Slimane Hammoudi, Alfredo Cuzzocrea, Abdelkrim Meziane, Amel Benna

ENCODE: Ensemble Contextual Bandits in Big Data Settings - A Case Study in E-Commerce Dynamic Pricing 5372
Srividhya Sethuraman, G Uma Maheswari, Siddhesh Thombre, Sunny Kumar, Vikash Patel, Sharadha Ramanan

Correlation and Unintended Biases on Univariate and Multivariate Decision Trees 5382
Mattia Setzu, Salvatore Ruggieri

Distributed Concept Drift Detection for Efficient Model Adaptation with Big Data Streams 5392
Ian Whitehouse, Rodrigo Yepez-Lopez, Roberto Corizzo

Task Unloading Algorithm for Mobile Edge Computing Based on Graph Neural Network 5398
Zhang Yuze, Xia Geming, Li Hongcheng, Li Hongfeng, Song Hui

Quantum Cryptography for Enhanced Network Security: A Comprehensive Survey of Research, Developments, and Future Directions 5408
Mst Shapna Akter, Juanjose Rodriguez-Cardenas, Hossain Shahriar, Alfredo Cuzzocrea, Fan Wu

A Trustable LSTM-Autoencoder Network for Cyberbullying Detection on Social Media Using Synthetic Data 5418
Mst Shapna Akter, Hossain Shahriar, Alfredo Cuzzocrea, Fan Wu, Juanjose Rodriguez-Cardenas

Adversarial Data-Augmented Resilient Intrusion Detection System for Unmanned Aerial Vehicles 5428
Muneeba Asif, Mohammad Ashiqur Rahman, Kemal Akkaya, Hossain Shahriar, Alfredo Cuzzocrea

Safe Data Transmission in IoT System for Air Pollution Monitoring 5438
Janusz Bobulski, Sabina Szymoniak, Kamila Pasternak

Private Continuous Survival Analysis with Distributed Multi-Site Data 5444
Luca Bonomi, Marilyn Lionts, Liyue Fan

Towards an Evolution in the Characterization of the Risk of Re-Identification of Medical Images 5454
Antoine Boutet, Carole Frindel, Mohamed Maouche

Randomized Response Has No Disparate Impact on Model Accuracy 5460
Alycia N. Carey, Karuna Bhaila, Xintao Wu

BigData Fusion for Trajectory Prediction of Multi-Sensor Surveillance Information Systems 5466
G. Cascavilla, A. Cuzzocrea, D. De Pascale, M. Omidbakhsh, D. A. Tamburri

ANTI-Disinformation: An Adversarial Attack and Defense Network Towards Improved Robustness for Disinformation Detection on Social Media 5476
Kuan-Chun Chen, Chih-Yao Chen, Cheng-Te Li

Matching Noisy Keys for Obfuscation 5485
Charlie Dickens, Eric Bax

Blockchain-Based Decentralized Verifiable Credentials: Leveraging Smart Contracts for Privacy-Preserving Authentication Mechanisms to Enhance Data Security in Scientific Data Access 5493
Md Jobair Hossain Faruk, Jim Basney, Jerry Q. Cheng

Multilingual Hate Speech Detection: Comparison of Transfer Learning Methods to Classify German, Italian, and Spanish Posts 5503
Jan Fillies, Michael Peter Hoffmann, Adrian Paschke

Targeted Ads Analysis: What Are the Most Targeted Personas? 5512
Hiroaki Kikuchi, Ayaka Aoyama

Privacy-Preserving Data Sharing in Agriculture: Enforcing Policy Rules for Secure and Confidential Data Synthesis 5519
Anantaa Kotal, Lavanya Elluri, Deepti Gupta, Varun Mandalapu, Anupam Joshi

Protecting Multiple Sensitive Attributes in Synthetic Micro-Data 5529
Nina Niederhametner, Rudolf Mayer

Ethical Hacking in Digital Forensics 5539
Damilola Oladimeji, Laura Garland, Jayanthi Ramamoorthy, Bing Zhou

Differentially Private Federated Learning: Privacy and Utility Analysis of Output Perturbation and DP-SGD 5549
Anastasia Pustozerova, Jan Baumbach, Rudolf Mayer

FedCIO: Efficient Exact Federated Unlearning with Clustering, Isolation, and One-Shot Aggregation 5559
Hongyu Qiu, Yongwei Wang, Yonghui Xu, Lizhen Cui, Zhiqi Shen

A Comparison of Federated Aggregation Strategies and Architectures for Next-Word Prediction 5569
Yana Sakhnovych, Richard Röttger, Rudolf Mayer

Achieving Privacy and Tracing Unauthorised Usage: Anonymisation-Based Fingerprinting of Private Data 5578
Tanja Šarcevic, Rudolf Mayer, Philipp Adler

Enhancing Object Detection in YouTube Thumbnails Forensics with YOLOv4 5588
Shahrzad Sayyafzadeh, Hongmei Chi, Shuyuan Mary Ho, Idongesit Mkpong-Ruffin

Sorting the Digital Stream: Big Data-Driven Insights into Email Classification for Spam and Ham Detection 5598
Syed Attique Shah, Emil Anthony Arputham, Awais Ahmed, Mohamed Ben Farah, Attal Shah, Abdul Aziz

Differentially Private Frequency Tables Based on Random Sampling 5608
Takumi Sugiyama, Kazuhiro Minami

[Vision Paper] Privacy-Preserving Data Integration 5614
Lisa Trigiante, Domenico Beneventano, Sonia Bergamaschi

Analyzing Code Text Strings for Code Evaluation 5619
Tianyu Wang, Zhixiong Chen

Topic-Based Search: Dataset Search Without Metadata and Users' Knowledge About Data 5629
Yukihisa Fujita, Teruaki Hayashi, Masahiro Kuwahara

Exploring the Fundamental Units of Semantic Representation of Data Using Heterogeneous Variable Network in Data Ecosystems 5639
Teruaki Hayashi, Yukihisa Fujita, Masahiro Kuwahara

Multi-View Graph Convolution for Participant Recommendation 5647
Xiaolong Liu, Liangwei Yang, Chen Wang, Mingdai Yang, Zhiwei Liu, Philip S. Yu

Variable-Based Learning Considering Topic Specificity in Heterogeneous Data Clustering Tasks 5657
Kosuke Manabe, Yukihisa Fujita, Masahiro Kuwahara, Teruaki Hayashi

Resale-Proof Trades of Data Under Budget Constraints: A Subject Experiment 5665
Kazuhito Ogawa, Naoki Watanabe

Indexing and Visualization of Climate Change Narratives Using BERT and Causal Extraction 5674
Hiroki Sakaji, Noriyasu Kaneda

From Base to Conversational: Japanese Instruction Dataset and Tuning Large Language Models 5684
Masahiro Suzuki, Masanori Hirano, Hiroki Sakaji

Harnessing Behavioral Traits to Enhance Financial Stock Recommender Systems: Tackling the User Cold Start Problem 5694
Takehiro Takayanagi, Kiyoshi Izumi

Religion and Spirituality on Social Media in the Aftermath of the Global Pandemic 5704
Olanrewaju Tahir Aduragba, Jialin Yu, Alexandra I. Cristea

Non-Uniform Sampling Methods for Large Itemset Mining 5714
Zahra Moteshaker Arani, Mostafa Haghir Chehreghani, Morteza Haghir Chehreghani

Domain-Enhanced Attention Enabled Deep Network for Behavioral Health Identification from 911 Narratives 5723
Abm Adnan Azmee, Martin Brown, Md Abdullah Al Hafiz Khan, Dominic Thomas, Yong Pei, Monica Nandan

Frequent Subgraph Mining in Dynamic Databases 5733
Zhaoming Chen, Xinyang Chen, Guoting Chen, Wensheng Gan

One-GPT: A One-Class Deep Fusion Model for Machine-Generated Text Detection 5743
Roberto Corizzo, Sebastian Leal-Arenas

Identifying Suitable Attributes for Record Linkage Using Association Analysis 5753
Nachiket Deo, Sanguthevar Rajasekaran, Rany Kamel

DG-GLNet: A Graph Convolutional Network for Prediction of Time Sequence with Topological Structure 5761
Guohui Ding, Yankai Wang, Yuqi Liu, Chao Jiang, Lulu Sun

Self-Supervised Federated Learning for Anomaly Detection 5770
Tze-Qian Eng, Hsing-Kuo Pao, Chi-Chen Liao

Fake News Detection Via Sentiment Neutralization 5780
Hanjuan Huang, Hsuan-Ting Peng, Hsing-Kuo Pao

SBTREC - A Transformer Framework for Personalized Tour Recommendation Problem with Sentiment Analysis 5790
Ngai Lam Ho, Roy Ka-Wei Lee, Kwan Hui Lim

Deciphering Student Coding Behavior: Interpretable Keystroke Features and Ensemble Strategies for Grade Prediction 5799
Muhammad Fawad Akbar Khan, John Edwards, Paul Bodily, Hamid Karimi

An Analysis of the Dynamics of Ties on Twitter 5809
Kiana Kheiri, Muhammad Fawad Akbar Khan, Tyler Derr, Hamid Karimi

IndexAI: AI Based Index Selection for NoSQL Databases 5818
Mohammad Mahdi Khosravi, Pinar Karagoz, Ismail Hakki Toroslu

Multi-Label Co-Training Using a Graph Neural Network. 5826
Ragini Kihlman, Maria Fasli

A Study to Compare YOLO Family Models for Metal Surface Anomaly Detection 5835
Abdüsselam Koç, Özlem Albayrak, Yasar Kurt, Perin Ünal

A Novel Method to Enable Transfer Learning of Structural Graph Representations 5842
Janet Layne, Edoardo Serra

Vision Paper: Advancing of AI Explainability for the Use of ChatGPT in Government Agencies – Proposal of a 4-Step Framework 5852
Alvina Hui Shan Lee, Venky Shankararaman, Eng Lieh Ouh

Multi-Source Selective Transfer Learning for Fake News Detection in New Event 5857
Ke Li, Bin Guo, Siyuan Ren, Yasan Ding, Zhiwen Yu

Process Mining Approach for Multi-Cloud SLA Reporting 5867
Jeremy Mechouche, Mohamed Sellami, Zakaria Maamar, Roua Touihri, Walid Gaaloul

Efficient Dynamic Federated Learning for Imbalanced Data 5877
Kasim Oztoprak, Reza Hassanpour

Understanding Divergent Framing of the Supreme Court Controversies: Social Media Vs. News Outlets 5880
Jinsheng Pan, Zichen Wang, Weihong Qi, Hanjia Lyu, Jiebo Luo

A Posteriori Trading-Inspired Model-Free Time Series Segmentation 5888
Mogens Graf Plessen

Investigating the Effectiveness of Deep Learning and CFA Interpolation Based Classifiers on Identifying AIGC 5897
Michael Reidy, Henry Mallon, Jiebo Luo

Batch Mode Deep Active Learning for Regression on Graph Data 5904
Peter Samoaa, Linus Aronsson, Philipp Leitner, Morteza Haghir Chehreghani

Seasonal-Trend Time Series Decomposition on Graphics Processing Units 5914
Dmitry Serykh, Stefan Oehmcke, Cosmin Oancea, Dainius Masiliunas, Jan Verbesselt, Yan Cheng, Stéphanie Horion, Fabian Gieseke, Nikolaj Hinnerskov

On Chatbots for Visual Exploratory Data Analysis 5924
Brodrick Stigall, Ryan Rossi, Jane Hoffswell, Xiang Chen, Shunan Guo, Fan Du, Eunyee Koh, Kelly Caine

A Rewiring Contrastive Patch PerformerMixer Framework for Graph Representation Learning 5930
Zhongtian Sun, Anoushka Harit, Alexandra I. Cristea, Jingyun Wang, Pietro Lio

Combining Transformers and Tree-Based Decoders for Solving Math Word Problems 5940
Jenq-Haur Wang, Yi-Fang Chen

Intelligent Attack Detection in ROS-Based Systems 5946
Hande Çavsi Zaim, Esra Nergis Yolaçan, Uraz Yavanoglu

A Fine-Grained Analysis of Public Opinion Toward Chinese Technology Companies on Reddit 5951
Enting Zhou, Yurong Liu, Hanjia Lyu, Jiebo Luo

TreeBERT: Advanced Representation Learning for Relation Extraction 5960
Shashank Hebbar, Ying Xie, Jiang Zhong

Targeted Querying of Closed High-Utility Itemsets 5967
Shan Huang, Wensheng Gan, Jinbao Miao

USER: Towards High-Utility Sequential Rules with Repetitive Items 5977
Hong Lin, Wensheng Gan, Gengsen Huang, Philip S. Yu

LM-DiskANN: Low Memory Footprint in Disk-Native Dynamic Graph-Based ANN Indexing 5987
Yu Pan, Jianxin Sun, Hongfeng Yu

Explainability of Leverage Points Exploration for Customer Churn Prediction 5997
Danny Y. C. Wang, Lars Arne Jordanger, Jerry Chun-Wei Lin

Accelerating Web-Based Graph Visualization with Pixel-Based Edge Bundling 6005
Jieting Wu, Jianxin Sun, Xinyan Xie, Tian Gao, Yu Pan, Hongfeng Yu

Mining Rare Utility Patterns Within Target Items 6015
Yixin Xu, Cuiwei Peng, Jiahui Chen, Wensheng Gan, Shicheng Wan

When Contrastive Learning Meets Graph Unlearning: Graph Contrastive Unlearning for Link Prediction 6025
Tzu-Hsuan Yang, Cheng-Te Li

Data Streaming in Concept-Dependent Granulation 6033
Radoslaw Cybulski, Piotr Artiemjew

Time and Space Complexity of Deterministic and Nondeterministic Decision Trees: Local Approach 6040
Kerven Durdymyradov, Mikhail Moshkov

Effective Face Inpainting by Conditional Generative Adversarial Network 6046
Tzung-Pei Hong, Jin-Hang Wu, Ja-Hwung Su, Tang-Kai Yin

Context-Based Fact-Checking Using Knowledge Graph 6051
Hyun Jung Lee, Mye Sohn

Analysis of Medical Incident Reports Using Text Mining 6057
Tomohiro Kimura, Shusaku Tsumoto, Shoji Hirano

Using Hospital Information System Data to Estimate Nursing Care Needs 6063
Hiroyuki Ikedo, Tomohiro Kimura, Shusaku Tsumoto, Shoji Hirano

Breaking the U: Asymmetric U-Net for Object Recognition in Muon Tomography☆ 6069
Jannic Cutura, Stefan Pasch

Iterative Deep Learning for Muon Scattering Tomography 6076
Dymitr Ruta, Robert Ruta

IEEE BigData Cup 2023 Report: Object Recognition with Muon Tomography Using Cosmic Rays 6084
Mateusz Wnuk, Jan Dziuba, Andrzej Janusz, Dominik Slezak

IEEE BigData 2023 Keystroke Verification Challenge (KVC) 6092
Giuseppe Stragapede, Ruben Vera-Rodriguez, Ruben Tolosana, Aythami Morales, Ivan Deandres-Tame, Naser Damer, Julian Fierrez, Javier-Ortega Garcia, Nahuel Gonzalez, Andrei Shadrikov, Dmitrii Gordin, Leon Schmitt, Daniel Wimmer, Christoph Großmann, Joerdis Krieger, Florian Heinz, Ron Krestel, Christoffer Mayer, Simon Haberl, Helena Gschrey, Yosuke Yamagishi, Sanjay Saha, Sanka Rasnayaka, Sandareka Wickramanayake, Terence Sim, Weronika Gutfeter, Adam Baran, Mateusz Krzysztón, Przemyslaw Jaskóla

Enhancing Digital Learning: A Usability Evaluation of the Canvas LMS at the University of North Texas a Pilot Study at the University of North Texas 6101
Shabbab Algamdi, Stephanie Ludi

Brain Stroke Prediction Using Machine Learning Techniques 6104
Ibrahim Almubark

Towards Large Language Model Organization: A Case Study on Abstractive Summarization 6109
Krisztián Boros, Masafumi Oyamada

Assessing Regional Disparities in Human Development and Multidimensional Poverty: A Satellite Imagery and Machine Learning Approach 6113
Sukbeom Chang, Youeel Sadek Kamal Abdelnour, Ivis María Companioni Cardoso, Jeongho Kim, Seon Tae Kim, Joo Hee Oh

Master-ASL: Sign Language Learning and Assessment System 6116
Sanjusha Cheemakurthi, Min Chen

An Augmentation-Agnostic Semantic Preserving Technique for Data Generation 6119
Jang-Ho Choi, Moonyoung Chung, Jiyong Kim

Text Mining Using Clinical Terms in Electronic Records of Annual Falls of Patients in Home Community Care 6122
Dillon Chrimes, Emile Keruzore

Big Data Usability Text Mining of Publicly Available YouTube Electronic Health Record (EHR) Tutorials 6125
Dillon Chrimes, Ivan Tang

Comparison of MIMIC-III and MIMIC-IV for Big Data Analytics of Health Informatics 6128
Dillon Chrimes, Chanhee Kim

Exploring Power Outage Prediction Using Weather and Socioeconomic Data in the Southeastern Part of the United States 6138
Ethan Cruz, Thilanka Munasinghe

Investigating Data Reusability in Density Functional Theory Studies 6143
Rob Fleur, Addy Ireland, Xintong Zhao, Scott McClellan, Eric Paltoo, Tianyu Su, Channyung Lee, Yuan An, Xiaohua Hu, Elif Ertekin, Jane Greenberg

Exploring AndroidManifest.xml for Automated Android Apps Classification 6145
Elena Flondor

Introducing the Partial Clustering Problem 6148
Byron J. Gao

Multi-View Latent Diffusion 6152
G. Di Giacomo, G. Franzese, T. Cerquitelli, C. F. Chiasserini, P. Michiardi

Enhanced Arrhythmia Classification Through Feature Engineering and Hierarchical Modeling 6155
Fahad Hakami, Lin-Ching Chang

Improved Dynamic Coupled Graph Convolutional Recurrent Networks for Traffic Forecasting 6159
Seunghoon Han, Sungsoo Kim, Sungsu Lim

Image Data Augmentation and Detection Study for Pothole Detection Algorithm 6162
Jisoo Hong, Youngjin Jung, Kang Sung Woo

Complex Networks Exploration with Triangles 6165
Shahadat Hossain, Raheem Mir, Emam Hossain

Design and Implementation of Blockchain Data Analysis Platform 6169
Miyoung Jang, Dong-Myung Sul, Jongho Won, Ji-Yong Kim

Voice of Employee: Impact of Online Reviews on Company and Job Seeker Matching Performance 6172
Haeun Jang, Sanghee Kim, Jongseok Jeon, Joohee Oh

Pineapple Quality Classification in a Multimodal Audio-Visual Dataset 6175
Yi-Lu Jiang, Wen-Chang Chang, Chih-Yi Chiu

Combining Wavelet and MFCC Features for Emotion Recognition in Conversation 6178
Geuna Jo, Hyuk Lim, Seunghyun Yoon

Ammunition Management in the AI Era: Towards CBM+ and Shelf-Life Analysis 6180
Youngjin Jung, Jisoo Hong, Solip Kim, Kang Sung Woo

Interpreting and Understanding the Image Using Deep Learning and Imprecise Information from the Monitoring System 6183
Mariusz Kubanek, Janusz Bobulski, Lukasz Karbowiak

Fast Graph Learning for Traffic Prediction 6186
Sungsoo Kim, Moonyoung Chung

A Computational Method of Identifying False Positives in the Variant Calling 6189
Sunhee Kim, Sang-Ho Chu, Yong-Jin Park, Chang-Yong Lee

Performance Evaluation of Data Imputation Methods for Graph Deep Learning-Based Traffic Prediction 6192
Jeongseon Kim, Sungsoo Kim, Sungsu Lim

Bi-Directional Face Tracking for Precise Facial De-Identification in Video Data 6195
Suha Kim, Hyuk Lim, Seunghyun Yoon

Detecting Greenwashing in Sustainability Disclosures: A Prediction Model for KOSPI 200 Enterprises Using ESG-BERT 6198
Seonu Kim, Yoel Shin, Seongwoo Park, Semakula Joel, Seon Tae Kim, Joo Hee Oh

A Data-Driven Approach to Predict Social Impact of Rural Tourism: Insights from UNWTO's Video Campaigns 6201
Kamuna Kipa, Nigel Kari Totona, Eunbi Cho, Jeongho Kim, Seon Tae Kim, Joo Hee Oh

Optimization of Public Bus Route in a Smart City Utilizing Real-Time Human Location Data Analysis 6204
Ryosuke Kobayashi, Nobuo Shigeta, Nobuyuki Saji, Rie Shigetomi Yamaguchi

Temporal Aspects of Cyber Threat Intelligence 6207
Angel Kodituwakku, Clark Xu, Daniel Rogers, David K. Ahn, Errin W. Fulp

Graph Neural Networks for Entity Matching 6212
Evgeny Krivosheev, Katsiaryna Mirylenka, Mattia Atzeni, Paolo Scotton

Energy BigData Platform for Residential Heat Information 6215
Tai-Yeon Ku, Wan-Ki Park, Hoon Choi

Utilizing Text Mining and GIS Analysis on the Vulnerable Disaster Type of the Elderly Living Alone 6218
Jinhyeok Lee, Hyunjung Kim

Threat Hunting in Internet of Things Networks with Bio-Inspired Models 6221
Jia Mei Grace Lim, Olakunle Olayinka

Feature Importances for Predicting Future Performance of Professional Soccer Player 6224
Gawon Lim

Big Data and Data Visualization Challenges 6227
Manoj Muniswamaiah, Tilak Agerwala, Charles C. Tappert

Comparison of SQL, NoSQL, and NewSQL Database Technologies 6230
Manoj Muniswamaiah, Tilak Agerwala, Charles C. Tappert

IoT-Based Big Data Storage Systems Challenges 6233
Manoj Muniswamaiah, Tilak Agerwala, Charles C. Tappert

Mitigating Annotation Artifacts in Physical Commonsense Reasoning Benchmark 6236
Wing-Lam Mok, Sungho Kim

Learning for Spatio-Temporal and Relational Data 6239
Kihyuk Nam, Taewhi Lee, Sung-Soo Kim, Insik Shin

Privacy-Preserving Approximate Query Processing with Differentially Private Generative Models 6242
Jiwon Ock, Taewhi Lee, Seongmin Kim

Evaluating Speech Emotion Recognition Through the Lens of CNN & LSTM Deep Learning Models 6245
Daniel F. O. Onah, Asia Ibrahim

AI Composite Sensor Using BigData for Energy Management System 6248
Wan-Ki Park, Tai-Yeon Ku

Pangolin: Incast Avoidance in Large Datacenter Networks 6251
Sae Hyong Park, Namseok Ko

Dividing and Merging for Imbalanced Long-Tail Datasets 6254
Minseo Park, Hyuk Lim, Seunghyun Yoon

Predicting Patent Transfer in the Manufacturing Industry: A Machine Learning Model for Patent Analytics Using BERT 6257
Hyeonmin Park, Jeongho Kim, Eunbi Cho, Joo Hee Oh

Compulsive Bingers Or Faithful Fans: Binge-Watching Behaviour Analysis Using Viewing Data and Twitter Discourse for 'The Traitors US' 6260
Joanne Parkes, Andrew Smith, Steve Benford, Michael Evans, Lianne Kerlin

AI in Archival Holdings and Digital History: Interrelational, Dialogical and Cooperative Interdisciplinary Partnerships 6263
Pepita Raventós Pajares, Meritxell Simon Martin, Josep Ballber Botey, Celio Hernández Tornero, Roberto García González, Glòria Jové Monclús

A Study on Watching Support of Werewolf Game Using Log Data Analyzed with Machine Learning 6265
Ryo Saito, Keita Ushida, Tatsuya Arakawa

Towards Automated Clustering 6268
Georg Stefan Schlake, Christian Beecks

Towards Retention Analysis in NFT-Based Blockchain Games 6271
Rafet Sifa, Fabian Hadiji

A Comparative Analysis of Data Augmentation Approaches for Improved Minority Behavior Detection in Digital Games 6275
Rafet Sifa, Edwin Yang

The Effect of Tweets on the Traded Volume of Crypto-Coins 6279
Aakhya Singh, Chirag Bhatia, Sohom Ghosh

Towards a Memory-Adaptive Hybrid Hash Join Design 6283
Giulliano Silva Zanotti Siviero, Shiva Jahangiri

Recurrent Traffic Demand Generation Using Urban Traffic Simulation with Cell-Based Behavior Model and Real Traffic Data 6286
Hyewon Song, Moonyoung Chung

Development of AI-Based Intrusion Detection System with Real-Time Flow Feature Extraction 6289
Giwon Sur, Hyejin Kim, Seunghyun Yoon, Hyuk Lim

Functional Arm Movement Classification in Stroke Survivors Using Deep Learning with Accelerometry Data 6292
Tan Tran, Lin-Ching Chang, Peter S. Lum

How Wildfire Wreaked Havoc in California: Impact of Evacuation on Residents? 6295
Vasuda Trehan

AI-Powered Archives: Revolutionizing Information Access for the Future 6298
Vasuda Trehan

Classification of Cognitive-Aesthetic Responses Using Machine Learning and EEG Data 6301
Edward Trudeau, Lin-Ching Chang, Yoshio Nakamura, Mohamad Koubeissi, Julio Bermudez

Proactive Resource Management to Optimize Distributed Workflow Executions 6305
Joel Witzke, Florian Schintke, Ansgar Lößer, Björn Scheuermann

Why is the User Interface a Dark Pattern?: Explainable Auto-Detection and Its Analysis 6308
Yuki Yada, Tsuneo Matsumoto, Fuyuko Kido, Hayato Yamana

Comprehensive Cough Data Analysis on CODA TB 6311
Jyoti Yadav, Aparna S. Varde, Lei Xie

Data Mining of Electronic Structure of Proteins and Quantum Evolution 6314
Masanori Yamanaka

An Alternative Illustration to Generalize Sprague-Grundy Theorem 6317
Dongho Yang, Sangwhan Cha

When LLM Meets Material Science: An Investigation on MOF Synthesis Labeling 6320
Xintong Zhao, Kyle Langlois, Jacob Furst, Scott McClellan, Rob Fleur, Yuan An, Xiaohua Hu, Fernando Uribe-Romo, Diego Gualdron, Jane Greenberg

Author Index