The 22nd IEEE International Conference on Data Mining

https://icdm22.cse.usf.edu/

ICDM-2022

Nov. 28 – Dec. 1
Orlando, Florida, USA

Sponsors

The 22nd IEEE International Conference on Data Mining

IEEE ICDM 2022 Program Schedule (Tentative)

November 28 - December 1, 2022

Keynote Lecture: **60 minutes** (about 45 minutes for talk and 15 minutes for Q and A)
Main conference regular paper: **20 minutes** (about 15 minutes for talk and 5 minutes for Q and A)
Main conference short paper: **15 minutes** (about 10 minutes for talk and 5 minutes for Q and A)

All conference activities are based on **USA Eastern Standard time.**

Day 1: Monday, November 28, 2022		
Time	**Logistics**	**Locatins**
08:30– 09:00	Continental Breakfast	Florida Foyers
10:30 – 11:00	Coffee Break	Florida Foyers
15:00 – 15:30	Coffee Break	Florida Foyers
Time	**Workshops**	**Chairs**
9:00-18:00	Machine Learning on Graphs (MLoG)	
	Incremental classification and clustering, conceptdrift, novelty detection, active learning in big/fast data context (IncrLearn)	
	International Workshop on Spatial and Spatio-Temporal Data Mining (SSTDM)	
	Optimization Based Techniques for Emerging Data Mining (OEDM)	
	International Workshop on AI for Nudging and Personalization (WAIN)	
	Sentiment Elicitation from Natural Text for Information Retrieval and Extraction (SENTIRE)	
	Workshop on Utility-Driven Mining and Learning (UDML)	
	Data Mining in Earth System Science (DMESS)	
	International Sustainable AI Workshop (ISAW)	
	Data Mining for Service (DMS)	
Time	**Workshops**	**Chairs**
	International Workshop on Knowledge Graphs (KG)	
	Workshop on Data Mining in Biomedical Informatics and Healthcare (DMBIH)	
	Machine Learning for Cybersecurity (MLC)	

	Workshop on AI for Aging, Rehabilitation and Intelligent Assisted Living (ARIAL)
	Urban Internet-of-Things Intelligence (UNIT)
	Workshop on High Dimensional Data Mining (HDM)
	Multi-view Representation Learning (MRL)
	Advanced Neural Algorithms and Theories for Recommender Systems (NeuRec)
	Data Mining in Learning Science (DMLS)
	Deep Learning and Clustering (DLC)
	Machine Learning on Higher-Order Structured data (ML-HOS)
	Foundation Models for Vision and Language (FOMO-VL)
	Workshop of Social Data Mining in the Post-pandemic Era (SocDM)
	Evolutionary Data Mining and Machine Learning (EDMML)

Day 2: Tuesday, November 29, 2022		
8: 45 – 9:00 Location: Florida 4	**Opening and Welcome** Conf Chairs, PC Chairs, Local Chair	
9:00 – 10:00 Location: Florida 4	**Keynote**: Dr. Eric P. Xing **Session Chair:**	
10: 00 - 10: 30 Florida Foyers	Coffee Break	
Time	**Sessions/Forums**	**Session Chair**
10:30-12:00	A1	
	A2	
	A3	
	T1	
	F1	
12:00 - 13:00 Florida 1-3	Lunch	
13:00-15:00	A4	
	A5	
	A6	
	T1	
	T2	

15:00-15:30 Florida Foyers	Coffee Break	
15:30-18:00	A7	
	A8	
	A9	
	T2	
18:00 – 19:00 Location: TBA	Reception	

Day 3: Wednesday, November 30, 2022		
9:00 – 10:00 Location: Florida 4	**Keynote**: Dr. Vipin Kumar **Session Chair:**	
10: 00 - 10: 30 Florida Foyers	Coffee Break	
Time	**Sessions/Forums**	**Session Chair**
10:30-12:00	B1	
	B2	
	B3	
	T3	
	NSF Session	
12:00 - 13:00 Florida 1-3	Lunch	
13:00-15:00	B4	
	B5	
	B6	
	T3	
	T4	
	Panel Discussion	
15:00-15:30 Florida Foyers	Coffee Break	
15:30-18:00	B7	
	B8	
	B9	
	T4	
19:00 – 21:00 Location: Florida 4	Banquet	

Day 4: Tursday, December 1, 2022		
9:00 – 10:00 Location: Florida 4	**Keynote**: Dr. Cynthia D. Rudin **Session Chair:**	
10: 00 - 10: 30 Florida Foyers	Coffee Break	
Time	**Sessions/Forums**	**Session Chair**
13:00-12:00	C1	
	C2	
	C3	
	F2	
12:00 - 13:00 Florida 1-3	Lunch	
13:00-15:00	C4	
	C5	
	C6	
15:00-15:30 Florida Foyers	Coffee Break	
15:30-18:00	C7	
	C8	
	C9	
18:00 – 18:00 Location: Florida 4	Concluding	

Keynote Lectures

Keynote 1: On Learning, Meta-Learning, and "Lego-Learning" -- theory, system, and engineering

Speaker:

Dr. Eric Poe Xing, Professor, Machine Learning Department, Carnegie Mellon University, USA

Abstract:

Software systems for complex tasks - such as controlling manufacturing processes in real-time; or writing radiological case reports within a clinical workflow – are becoming increasingly sophisticated and consist of a large number of data, model, algorithm, and system elements and modules. Traditional benchmark/leaderboard-driven bespoke approaches in the Machine Learning community are not suited to meet the highly demanding industrial standards beyond algorithmic performance, such as cost-effectiveness, safety, scalability, and automatability, typically expected in production systems. In this talk, I discuss some technical issues toward addressing these challenges: 1) a theoretical framework for panoramic learning with all experiences; 2) optimization methods to best the effort for learning under such a principled framework; 3) compositional strategies for building production-grade ML programs from standard parts. I will present our recent work toward developing a "Standard Model" for Learning that unifies different special-purpose machine learning paradigms and algorithms, then a Bayesian blackbox optimization approach to Meta Learning in the space of hyperparameters, model architectures, and system configurations, and finally principles and designs of standardized software Legos that facilitate cost-effective building, training, and tunning of practical ML pipelines and systems.

Short Bio:

Eric P. Xing is the President of the Mohamed bin Zayed University of Artificial Intelligence, a Professor of Computer Science at Carnegie Mellon University, and the Founder and Chairman of Petuum Inc., a 2018 World Economic Forum Technology Pioneer company that builds standardized artificial intelligence development platform and operating system for broad and general industrial AI applications. He completed his PhD in Computer Science at UC Berkeley. His main research interests are the development of machine learning and statistical methodology; and composable, automatic, and scalable computational systems, for solving problems involving automated learning, reasoning, and decision-making in artificial, biological, and social systems.

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

Speaker:

Dr. Vipin Kumar, Regents Professor and William Norris Chair in Large Scale Computing, Department of Computer Science and Engineering, University of Minnesota, USA

Abstract:

Process-based models of dynamical systems are often used to study engineering and environmental systems. Despite their extensive use, these models have several well-known limitations due to incomplete or inaccurate representations of the physical processes being modeled. There is a tremendous opportunity to systematically advance modeling in these domains by using state of the art machine learning (ML) methods that have already revolutionized computer vision and language translation. However, capturing this opportunity is contingent on a paradigm shift in data-intensive scientific discovery since the “black box” use of ML often leads to serious false discoveries in scientific applications. Because the hypothesis space of scientific applications is often complex and exponentially large, an uninformed data-driven search can easily select a highly complex model that is neither generalizable nor physically interpretable, resulting in the discovery of spurious relationships, predictors, and patterns. This problem becomes worse when there is a scarcity of labeled samples, which is quite common in science and engineering domains.

This talk makes the case that in real-world systems that are governed by physical processes, there is an opportunity to take advantage of fundamental physical principles to inform the search of a physically meaningful and accurate ML model. While this talk will illustrate the potential of the knowledge-guided machine learning (KGML) paradigm in the context of environmental problems (e.g., Fresh water science, Hydrology, Agronomy), the paradigm has the potential to greatly advance the pace of discovery in a diverse set of discipline where mechanistic models are used, e.g., climate science, weather forecasting, and pandemic management.

Research funded by NSF (Expeditions in Computing, BIGDATA, INFEWS, STC, GCR, and HDR programs), DARPA, ARPA-E, and USGS

Short Bio:

Vipin Kumar is a Regents Professor and holds William Norris Chair in the department of Computer Science and Engineering at the University of Minnesota. His research spans data mining, high-performance computing, and their applications in Climate/Ecosystems and health care. He also served as the Director of Army High Performance Computing Research Center (AHPCRC) from 1998 to 2005. He has authored over 400 research articles, and co-edited or coauthored 11 books including two widely used text books "Introduction to Parallel Computing", "Introduction to Data Mining", and a recent edited collection, "Knowledge Guided Machine Learning". Kumar's current major research focus is on knowledge-guided machine learning and its applications to understanding the impact of human induced changes on the Earth and its environment. Kumar's research on this topic is funded by NSF's BIGDATA, INFEWS, STC, GCR, and HDR programs, as well as ARPA-E, DARPA, and USGS. He has recently finished serving as the Lead PI of a 5-year, $10 Million project,Understanding Climate Change - A Data Driven Approach, funded by the NSF's Expeditions in Computing program. Kumar is a Fellow of the ACM, IEEE, AAAS, and SIAM. Kumar's foundational research in data mining and high performance computing has

been honored by the ACM SIGKDD 2012 Innovation Award, which is the highest award for technical excellence in the field of Knowledge Discovery and Data Mining (KDD), the 2016 IEEE Computer Society Sidney Fernbach Award, one of IEEE Computer Society's highest awards in high performance computing, and Test-of-time award from 2021 Supercomputing conference (SC21)

Keynote 3: Do Simpler Machine Learning Models Exist and How Can We Find Them?

Speaker:

Dr. Cynthia D. Rudin, Professor, Computer Science, Electrical and Computer Engineering, Statistical Science, and Biostatistics & Bioinformatics, Duke University, USA

Abstract:

While the trend in machine learning has tended towards building more complicated (black box) models, such models are not as useful for high stakes decisions - black box models have led to mistakes in bail and parole decisions in criminal justice, flawed models in healthcare, and inexplicable loan decisions in finance. Simpler, interpretable models would be better. Thus, we consider questions that diametrically oppose the trend in the field: for which types of datasets would we expect to get simpler models at the same level of accuracy as black box models? If such simpler-yet-accurate models exist, how can we use optimization to find these simpler models? In this talk, I present an easy calculation to check for the possibility of a simpler (yet accurate) model before computing one. This calculation indicates that simpler-but-accurate models do exist in practice more often than you might think. Also, some types of these simple models are (surprisingly) small enough that they can be memorized or printed on an index card.

Short Bio:

Cynthia D. Rudin is a professor of computer science, electrical and computer engineering, statistical science, mathematics, and biostatistics & bioinformatics at Duke University, and directs the Interpretable Machine Learning Lab. Previously, Prof. Rudin held positions at MIT, Columbia, and NYU. She holds an undergraduate degree from the University at Buffalo, and a PhD from Princeton University. She is the recipient of the 2022 Squirrel AI Award for Artificial Intelligence for the Benefit of Humanity from the Association for the Advancement of Artificial Intelligence (AAAI). This award is the most prestigious award in the field of artificial intelligence. Similar only to world-renowned recognitions, such as the Nobel Prize and the Turing Award, it carries a monetary reward at the million-dollar level. Prof. Rudin is also a three-time winner of the INFORMS Innovative Applications in Analytics Award, was named as one of the "Top 40 Under 40" by Poets and Quants in 2015, and was named by Businessinsider.com as one of the 12 most impressive professors at MIT in 2015, and is a 2022 Guggenheim Fellow. She is a fellow of the American Statistical Association, the Institute of Mathematical Statistics, and AAAI.

Prof. Rudin is past chair of both the INFORMS Data Mining Section and the Statistical Learning and Data Science Section of the American Statistical Association. She has also served on committees for DARPA, the National Institute of Justice, AAAI, and ACM SIGKDD. She has served on several committees for the National Academies of Sciences, Engineering and Medicine, including the Committee on Applied and Theoretical Statistics, the Committee on Law and Justice, and the Committee on Analytic Research Foundations for the Next-Generation

Conference Paper Presentations

Paper Sessions

[Sessions A1-A9 and B1-B9 are onsite presentations. Sessions C1-C9 are remote presentations]

A1	A1: Structure-aware Learning (3 full + 2 short)	
	DM207: Robust Structure-aware Semi-supervised Learning	Xu Chen (full)
	DM1013: How Out-of-Distribution Data Hurts Semi-Supervised Learning	Xujiang Zhao, Krishnateja Killamsetty, Rishabh Iyer, and Feng Chen (full)
	DM442: Conditional Independence Testing via Latent Representation Learning	Bao Duong and Thin Nguyen (full)
	DM806: Delta-Closure Structure for Studying Data Distribution	Aleksey Buzmakov, Sergei O. Kuznetsov, Tatiana Makhalova, and Amedeo Napoli (short)
	DM781: Dirichlet process mixture models for non-stationary data streams	Ioar Casado and Aritz Perez (short)
A2	A2: Graph Learning I (3 full + 2 short)	
	DM287: Generalized Few-Shot Node Classification	Zhe Xu, Kaize Ding, Yu-Xiong Wang, Huan Liu, and Hanghang Tong (full)
	DM395: Feature-Oriented Sampling for Fast and Scalable GNN Training	Xin Zhang, Yanyan Shen, and Lei Chen (full)
	DM639: Dimensionality Selection of Hyperbolic Graph Embeddings using Decomposed Normalized Maximum Likelihood Code-Length	Ryo Yuki, Yuichi Ike, and Kenji Yamanishi (full)
	DM662: Uncertainty Propagation in Node Classification	Zhao Xu, Carolin Lawrence, Ammar Shaker, and Raman Siarheyeu (short)
	DM385: Augmenting Knowledge Transfer across Graphs	Yuzhen Mao, Jianhui Sun, and Dawei Zhou (short)
A3	A3: Text and Set Learning (3 full + 2 short)	
	DM1052: Enhancing Robust Text Classification via Category Description	Xin Gao, Zhengye zhu, Xu Chu, Yasha Wang, Wenjie Ruan, and Junfeng Zhao (full)
	DM1098: HyperKGQA: Question Answering over Knowledge Graphs using Hyperbolic Representation Learning	Nadya Abdel Madjid, Ola El Khatib, Shuang Gao, and Djellel Difallah (ful)
	DM725: Text-enhanced Multi-Granularity Temporal Graph Learning for Event Prediction	Xiaoxue Han and Yue Ning (full)
	DM408: Set2Box: Similarity Preserving Representation Learning for Sets	Geon Lee, Chanyoung Park, and Kijung Shin (short)
	DM231: Deep Stable Representation Learning on Electronic Health Records	Yingtao Luo, Zhaocheng Liu, and Qiang Liu (short)
A4	A4: Graph Learning II (3 full + 4 short)	
	DM685: Unifying Graph Contrastive Learning with Flexible Contextual Scopes	Yizhen Zheng, Yu Zheng, Xiaofei Zhou, Chen Gong, Vincent CS Lee, and Shirui Pan (full)
	DM774: Backpropagation-free Graph Neural Networks	Luca Pasa, Nicolo Navarin, Wolfgang Erb, and Alessandro Sperduti (full)
	DM1030: GraphBERT: Bridging Graph and Text for Malicious Behavior Detection on Social Media	Jiele Wu, Chunhui Zhang, Zheyuan Liu, Erchi Zhang, Steven Wilson, and Chuxu Zhang (full)

	DM995: Sparsified Subgraph Memory for Continual Graph Representation Learning DM546: Jaccard Median for Ego-network Segmentation	Xikun Zhang, Dongjin Song, and Dacheng Tao (short) Haodi Zhong, Grigorios Loukides, Alessio Conte, and Solon Pissis (short)
	DM559: Diameter Minimization by Shortcutting with Degree Constraints	Florian Adriaens and Aristides Gionis (short)
	DM763: Near-Optimal Spectral Disease Mitigation in Healthcare Facilities	Masahiro Kiji, D. M. Hasibul Hasan, Alberto M. Segre, Sriram V. Pemmaraju, and Bijaya Adhikari (short)
A5	A5: Anomaly Detection (3 full + 4 short)	
	DM581: Unsupervised Deep Subgraph Anomaly Detection	Zheng Zhang and Liang Zhao (full)
	DM307: Toward Unsupervised Outlier Model Selection	Yue Zhao, Sean Zhang, and Leman Akoglu (full)
	DM1018: DAGAD: Data Augmentation for Graph Anomaly Detection	Fanzhen Liu, XiaoXiao Ma, Jia Wu, Jian Yang, Shan Xue, Amin Beheshti, Chuan Zhou, Hao Peng, Quan Z. Sheng, and Charu Aggarwal (full)
	DM1050: D.MCA: Outlier Detection with Explicit Micro-Cluster Assignments	Shuli Jiang, Robson Leonardo Ferreira Cordeiro, and Leman Akoglu (short)
	DM627: Post-Robustifying Deep Anomaly Detection Ensembles by Model Selection	Benedikt Boing, Simon Kluttermann, and Emmanuel Muller (short)
	DM959: DeepiForest: A Deep Anomaly Detection Framework with Hashing Based Isolation Forest	Haolong Xiang, Hongsheng Hu, and Xuyun Zhang (short)
	DM958: Detecting Irregular Network Activity with Adversarial Learning and Expert Feedback	Gopikrishna Rathinavel, Nikhil Muralidhar, Timothy OShea, and Naren Ramakrishnan (short)
A6	A6: Spatial-Temporal Learning (3 full + 4 short)	
	DM255: STrans-GAN: Spatially-Transferable Generative Adversarial Networks for Urban Traffic Estimation	Yingxue Zhang, Yanhua Li, Xun Zhou, Xiangnan Kong, and Jun Luo (full)
	DM257: Mest-GAN: Cross-City Urban Traffic Estimation with Meta Spatial-Temporal Generative Adversarial Networks	Yingxue Zhang, Yanhua Li, Xun Zhou, and Jun Luo (full)
	DM587: STORM-GAN: Spatio-Temporal Meta-GAN for Cross-City Estimation of Human Mobility Responses to COVID-19	Han Bao, Xun Zhou, Yiqun Xie, Yanhua Li, and Xiaowei Jia (full)
	DM696: Spatiotemporal Contextual Consistency Network for Precipitation Nowcasting	Xinyu Xiao, Qizhao Jin, Gaofeng Meng, Shiming Xiang, and Chunhong Pan (short)
	DM1072: Deep Spatial Domain Generalization	Dazhou Yu, Guangji Bai, Yun Li, and Liang Zhao (short)
	DM333: THOR: Self-Supervised Temporal Knowledge Graph Embedding via Three-Tower Graph Convolutional Networks	Yeon-Chang Lee, JaeHyun Lee, Dongwon Lee, and Sang-Wook Kim (short)
	DM624: THINK: Temporal Hypergraph Hyperbolic Network	Shivam Agarwal, Ramit Sawhney, Megh Thakkar, Preslav Nakov, Jiawei Han, and Tyler Derr (short)
A7	A7: Time Series, Trajectory, and Data Flow (3 full + 5 short)	
	DM285: Matrix Profile XXV: Introducing Novelets: A Primitive that Allows Online Detection of Emerging Behaviors in Time Series	Ryan Mercer and Eamonn Keogh (full)
	DM885: Neuro-symbolic Models for Interpretable Time Series Classification using Temporal Logic Description	Ruixuan Yan, Tengfei Ma, Achille Fokoue, Maria Chang, and Agung Julius (full)
	DM938: Class-Specific Explainability for Deep Time Series Classifiers	Ramesh Doddaiah, Prathyush Parvatharaju, Elke Rundensteiner, and Thomas Hartvigsen (full)
	DM467: Matrix Profile XXVI: Mplots: Scaling Time Series Similarity Matrices to Massive Data	Maryam Shahcheraghi, Ryan Mercer, Joao Manuel de Almeida Rodrigues, Audrey Der, Hugo Filipe Silveira Gamboa,

		Zachary Zimmerman, and Eamonn Keogh (short)
	DM810: Robust Time Series Chain Discovery with Incremental Nearest Neighbors	Li Zhang, Yan Zhu, Yifeng Gao, and Jessica Lin (short)
	DM1064: Self-explaining Hierarchical Model for Intraoperative Time Series	Dingwen Li, Bing Xue, Christopher King, Bradley Fritz, Michael Avidan, Joanna Abraham, and Chenyang Lu (short)
	DM289: Multi-view POI-level Cellular Trajectory Reconstruction for Digital Contact Tracing of Infectious Diseases	Dongmin Park, Junhyeok Kang, Hwanjun Song, Susik Yoon, and Jae-Gil Lee (short)
	DM909: Causal Effect Prediction with Flow-based Inference	Shaogang Ren, Dingcheng Li, and Ping Li (short)
A8	A8: Model Interpretabilit and Fairness (3 full + 5 short)	
	DM893: Concise and interpretable multi-label rule sets	Martino Ciaperoni, Aristides Gionis, and Han Xiao (full)
	DM737: ALICE and the Caterpillar: A More Descriptive Null Model for Assessing Data Mining Results	Giulia Preti, Gianmarco De Fracisci Morales, and Matteo Riondato (full)
	DM292: NeuCEPT: Learn Neural Networks's Mechanism via Critical Neurons with Precision Guarantee	Minh Vu, Truc Nguyen, and my thai (full)
	DM577: Unfair AI: It Isn't Just Biased Data	Chowdhury Mohammad Rakin Haider, Chris Clifton, and Yan Zhou (short)
	DM757: Fairness-Aware Graph Sampling for Network Analysis	Farzan Masrour, Francisco Santos, Pang-Ning Tan, and Abdol-Hossein Esfahanian (short)
	DM783: Toward Unified Data and Algorithm Fairness via Adversarial Data Augmentation and Adaptive Model Fine-tuning	Yanfu Zhang, Runxue Bao, Jian Pei, and Heng Huang (short)
	DM863: On Finding and Analyzing the Backbone of the k-Core Structure of a Graph	Ricky Laishram and Sucheta Soundarajan (short)
	DM931: Homology-Separating Triangulated Euler Characteristic Curve	Nicholas Malott, Robert Lewis, and Philip Wilsey (short)
A9	A9: Behavior Modeing and Recommendation (3 full + 5 short)	
	DM837: Widening the Time Horizon: Predicting the Long-Term Behavior of Chaotic Systems	Yong Zhuang, Matthew Almeida, Wei Ding, Patrick Flynn, Shafiqul Islam, and Ping Chen (full)
	DM272: Beyond Double Ascent via Recurrent Neural Tangent Kernel in Sequential Recommendation	Ruihong Qiu, Zi Huang, and Hongzhi Yin (full)
	DM900: Mitigating Popularity Bias in Recommendation with Unbalanced Interactions: A Gradient Perspective	Weijieying Ren, Lei Wang, Kunpeng Liu, Ruocheng Guo, LIM Ee Peng, and Yanjie Fu (full)
	DM383: Towards High-Order Complementary Recommendation via Logical Reasoning Network	Longfeng Wu, Yao Zhou, and Dawei Zhou (short)
	DM507: A Two-Stream Light Graph Convolution Network-based Latent Factor Model for Accurate Cloud Service QoS Estimation	Fanghui Bi, Tiantian He, and Xin Luo (short)
	DM549: Link Partitioning on Simplicial Complexes Using Higher-Order Laplacians	Xinyi Wu, Arnab Sarker, and Ali Jadbabaie (short)
	DM452: Knowledge-Guided Semantics Adjustment for Improved Few-Shot Classification	Guangtao Zheng and Aidong Zhang (short)
	DM434: Change Detection with Probabilistic Models on Persistence Diagrams	Kohei Ueda, Yuichi Ike, and Kenji Yamanishi (short)
B1	B1: Federated Learning and Ensembles (3 full + 2 short)	
	DM265: GraphFL: A Federated Learning Framework for Semi-Supervised Node Classification on Graphs	Binghui Wang, Ang Li, Meng Pang, Hai Li, and Yiran Chen (full)

	DM879: Federated Fingerprint Learning with Heterogeneous Architectures	Tianshi Che, Zijie Zhang, Yang Zhou, Xin Zhao, Ji Liu, Zhe Jiang, Da Yan, Ruoming Jin, and Dejing Dou (full)
	DM494: A Large-Scale Ensemble Learning Framework for Demand Forecasting	Young-Jin Park, Donghyun Kim, Frederic Odermatt, Juho Lee, and Kyung-Min Kim (full)
	DM1024: Personalized Federated Learning via Heterogeneous Modular Networks	Tianchun Wang, Wei Cheng, Dongsheng Luo, Wenchao Yu, Jingchao Ni, Liang Tong, Haifeng Chen, and Xiang Zhang (short)
	DM921: Boosting Object Detection Ensembles with Error Diversity	Ka Ho Chow and Ling Liu (short)
B2	B2: Factorization (3 full + 2 short)	
	DM593: Rethinking Symmetric Matrix Factorization: A More General and Better Clustering Perspective	Mengyuan Zhang and Kai Liu (full)
	DM743: Coresets remembered and items forgotten: submodular maximization with deletions	Guangyi Zhang, Nikolaj Tatti, and Aristides Gionis (full)
	DM987: TENALIGN: Joint Tensor Alignment and Coupled Factorization	Yunshu Wu, Uday Saini, Jia Chen, and Evangelos Papalexakis (full)
	DM689: Boolean Matrix Factorization for Data with Symmetric Variables	Jan Konecny and Martin Trnecka (short)
	DM728: Correntropy-Induced Tensor Learning for Multi-view Subspace Clustering	Yongyong Chen, Shuqin Wang, Jingyong Su, and Junxin Chen (short)
B3	B3: Graph Learning III (3 full +2 short)	
	DM517: GraphGDP: Generative Diffusion Processes for Permutation Invariant Graph Generation	Han Huang, Leilei Sun, Bowen Du, Yanjie Fu, and Weifeng Lv (full)
	DM979: HiGIL: Hierarchical Graph Inference Learning for Fact Checking	Qianren Mao, Yiming Wang, Chenghong Yang, Linfeng Du, Yuxin Ying, Jia Wu, Hao Peng, Jianxin Li, Lihong Wang, and Zheng Wang (full)
	DM379: Causality Enhanced Societal Event Forecasting with Heterogeneous Graph Learning	Songgaojun Deng, Huzefa Rangwala, and Yue Ning (ful)
	DM756: A Generic Graph Sparsification Framework using Deep Reinforcement Learning	Ryan Wickman, Xiaofei Zhang, and Weizi Li (short)
	DM956: HP-GMN: Graph Memory Networks for Heterophilous Graphs	Junjie Xu, Enyan Dai, Xiang Zhang, and Suhang Wang (short)
B4	B4: Online, Transfer, and Multi-Objective Learning (2 full + 5 short)	
	DM390: Efficient Navigation for Constrained Shortest Path with Adaptive Expansion Control	Wenwen Xia, Yuchen Li, Wentian Guo, and Shenghong Li (full)
	DM426: Projection Dual Averaging Based Second-order Online Learning	Zhong Chen, Huixin Zhan, Victor Sheng, Andrea Edwards, and Kun Zhang (full)
	DM569: DBHD: Density-based clustering for highly varying density	Walid Durani, Mautz Dominik, Claudia Plant, and Christian Bohm (short)
	DM564: Center-Based iPSC Colony Counting with Multi-Task Learning	Mirza Tanzim Sami, Da Yan, Bhadhan Roy Joy, Jalal Khalil, Ricardo Cevallos, Md Emon Hossain, Kejin Hu, and Yang Zhou (short)
	DM874: Multi-Objective Symbolic Regression for Data-Driven Scoring System Management	Davide Ferrari, Veronica Guidetti, and Federica Mandreoli (short)
	DM215: On Mean-Optimal Robust Linear Discriminant Analysis	Xiangyu Li and Hua Wang (short)

	DM373: Cost-Effective Transfer Learning for Data Streams	Ocean Wu, Yun Sing Koh, Gillian Dobbie, and Thomas Lacombe (short)
B5	B5: Security, Privacy, and Anomaly (3 full + 4 short)	
	DM551: Query-Efficient Target-Agnostic Black-Box Attack	Raha Moraffah and Huan Liu (full)
	DM746: Privacy-Preserving Split Learning via Patch Shuffling over Transformers	Dixi Yao, Liyao Xiang, Hengyuan Xu, Hangyu Ye, and Yingqi Chen (full)
	DM1026: Privacy-Preserved Neural Graph Similarity Learning	Yupeng Hou, Wayne Xin Zhao, Yaliang Li, and Ji-rong Wen (full)
	DM369: Few-shot Anomaly Detection and Classification Through Reinforced Data Selection	Xiao Han, Depeng Xu, Shuhan Yuan, and Xintao Wu (short)
	DM374: Making Reconstruction-based Method Great Again for Video Anomaly Detection	Yizhou Wang, Can Qin, Yue Bai, Yi Xu, Xu Ma, and Yun Fu (short)
	DM762: The Devil is in the Conflict: Disentangled Information Graph Neural Networks for Fraud Detection	Zhixun Li, Dingshuo Chen, Qiang Liu, and Shu Wu (short)
	DM325: MSDN: A Multi-Subspace Deviation Net for Anomaly Detection	Sinong Zhao, Zhaoyang Yu, Trent G. Marbach, Gang Wang, and Xiaoguang Liu (short)
B6	B6: Computational Models for Data Mining (3 full + 4 short)	
	DM497: Towards Scalable and Fast Distributionally Robust Optimization for Data-Driven Deep Learning	Xuli Shen, Xiaomei Wang, Qing Xu, Weifeng Ge, and Xiangyang Xue (full)
	DM381: Fast Stochastic Recursive Momentum Methods for Imbalanced Data Mining	Xidong Wu, Feihu Huang, and Heng Huang (full)
	DM827: Extremely Fast Hoeffding Adaptive Tree	Chaitanya Manapragada, Mahsa Salehi, and Geoffrey Webb (full)
	DM461: Communication-Efficient Adam-Type Algorithms for Distributed Data Mining	Wenhan Xian, Feihu Huang, and Heng Huang (short)
	DM344: Bayesian Deep Structured Semantic Model for Sub-Linear-Time Information Retrieval	Yuji Mo and Stephen Scott (short)
	DM217: Serialized Interacting Mixed Membership Stochastic Block Model	Gael Poux-Medard, Julien Velcin, and Sabine Loudcher (short)
	DM501: Estimating the Pruned Search Space Size of Subgroup Discovery	Lennart Purucker, Felix Stamm, Florian Lemmerich, and Joeran Beel (short)
B7	B7: Graph Learing IV (3 full + 5 short)	
	DM375: CC-GNN: A Community and Contraction-based Graph Neural Network	Zhiyuan Li, Xun Jian, Yue Wang, and Lei Chen (full)
	DM1005: Multi-Relational Graph Neural Architecture Search with Fine-grained Message Passing	Xin Zheng, Miao Zhang, Chunyang Chen, Chaojie Li, Chuan Zhou, and Shirui Pan (full)
	DM638: Revisiting Link Prediction on Heterogeneous Graphs with A Multi-view Perspective	Anasua Mitra, Priyesh Vijayan, Ranbir Sanasam, Diganta Goswami, Srinivasan Parthasarathy, and Balaraman Ravindran (full)
	DM567: DeepGAR: Deep Graph Learning for Analogical Reasoning	Chen Ling, Tanmoy Chowdhury, Junji Jiang, Junxiang Wang, Xuchao Zhang, Haifeng Chen, and Liang Zhao (short)
	DM873: Improving Graph Representation Learning with Distribution Preserving	Chengsheng Mao and Yuan Luo (short)
	DM372: Active Heterogeneous Graph Neural Networks with Per-step Meta-Q-Learning	Yuheng Zhang, Yinglong Xia, Yan Zhu, Yuejie Chi, Lei Ying, and Hanghang Tong (short)

	DM1003: Two Sides of the Same Coin: Heterophily and Oversmoothing in Graph Convolutional Neural Networks	Yujun Yan, Milad Hashemi, Kevin Swersky, Yaoqing Yang, and Danai Koutra (short)
	DM543: A Convex Formulation for Graph Convolutional Training: Two Layer Case	Naganand Yadati (short)
B8	B8: Health, Climate, Security Applications (3 full + 5 short)	
	DM439: MedSkim: Denoised Health Risk Prediction via Skimming Medical Claims Data	Suhan Cui, Junyu Luo, Muchao Ye, Jiaqi Wang, Ting Wang, and Fenglong Ma (full)
	DM784: Deep Geometrical Learning for Alzheimer's Disease Progression Modeling	Seungwoo Jeong, Wonsik Jung, Junghyo Sohn, and HEUNG-Il SUK (full)
	DM988: Meta-Transfer Learning: An application to Streamflow modelling in River-streams	Rahul Ghosh, Bangyan Li, Kshitij Tayal, Vipin Kumar, and Xiaowei Jia (full)
	DM637: Dengue Fever: From Extreme Climates to Outbreak Prediction	Thai Son Mai, Thu Ha Phi, Abdullahi Abubakar, Quoc Viet Hung Nguyen, Peter Kilpatrick, and Hans Vandierendonck (short)
	DM779: RoS-KD: A Robust Stochastic Knowledge Distillation Approach for Noisy Medical Imaging	Ajay Jaiswal, Kumar Ashutosh, Justin F Rousseau, Yifan Peng, Zhangyang Wang, and Ying Ding (short)
	DM890: Heterogeneous Information Enhanced Prerequisite Learning in Massive Open Online Courses	Tianqi Wang, Fenglong Ma, Yaqing Wang, and Jing Gao (short)
	DM965: DREAM: Domain Invariant and Contrastive Representation for Sleep Dynamics	Seungyeon Lee, Thai-Hoang Pham, and Ping Zhang (short)
	DM528: Adversarial Variational Modality Reconstruction and Regularization for Zero-Day Malware Variants Similarity Detection	Christopher Molloy, Jeremy Banks, Steven H. H. Ding, Philippe Charland, Andrew Walenstein, and Litao Li (short)
B9	B9: Deep Learning (1 full + 7 short)	
	DM378: ABN: Anti-Blur Neural Networks for Multi-Stage Deformable Image Registration	Yao Su, Xin Dai, Lifang He, and Xiangnan Kong (full)
	DM547: Pruning Adversarially Robust Neural Networks without Adversarial Examples	Tong Jian, Zifeng Wang, Jennifer Dy, Yanzhi Wang, and Stratis Ioannidis (short)
	DM999: Imbalanced Adversarial Training with Reweighting	Wentao Wang, Han Xu, Xiaorui Liu, Yaxin Li, Bhavani Thuraisingham, and Jiliang Tang (short)
	DM795: Deep Clustering with Consensus Representations	Lukas Miklautz, Martin Teuffenbach, Pascal Weber, Walid Durani, Rona Perjuci, Christian Bohm, and Claudia Plant (short)
	DM650: Verips: Verified Pseudo-label Selection for Deep Active Learning	Sandra Gilhuber, Philipp Jahn, Yunpu Ma, and Thomas Seidl (short)
	DM420: Robust Unsupervised Domain Adaptation from a Corrupted Source	Shuyang Yu, Zhuangdi Zhu, Boyang Liu, Anil Jain, and Jiayu Zhou (short)
	DM327: The Impact of Batch Learning in Stochastic Linear Bandits	Danil Provodin, Pratik Gajane, Mykola Pechenizkiy, and Maurits Kaptein (short)
	DM878: Reciprocity in Directed Hypergraphs: Measures, Findings, and Generators	Sunwoo Kim, Minyoung Choe, Jaemin Yoo, and Kijung Shin (short)
C1	C1: Deep Learning Applications (5 full) (Remote)	
	DM228: Learning Representations from Local to Global for Fine-grained Patient Similarity Measuring in Intensive Care Unit	Xianli Zhang, Buyue Qian, Yang Li, Zeyue Gao, Chong Guan, Renzhen Wang, Chen Li, and Yefeng Zheng (full)
	DM824: Self-Attention Gated Cognitive Diagnosis for Faster Adaptive Educational Assessments	Xiaohuan Pei, Shuo Yang, Jiajun Huang, and Chang Xu (full)

	DM354: Kernel-based Hybrid Interpretable Transformer for High-frequency Stock Movement Prediction	Fan Lin, Pengfei Li, Yuanguo Lin, Zhennan Chen, Huanyu You, and Shibo Feng (full)
	DM990: A Unified Guaranteed Impression Allocation Framework for Online Display Advertising	Hong Zhang, Lan Zhang, Ju Huang, Anran Li, Haoran Cheng, Dongbo Huang, and Lan Xu (full)
	DM963: FineFormer: Fine-Grained Adaptive Object Transformer for Image Captioning	Wang Bo, Zhang Zhao, Jicong Fan, Mingbo Zhao, and Mingliang Xu (full)
C2	C2: Learning and Fusion (5 full) (Remote)	
	DM310: Contrastive Code-Comment Pre-training	Xiaohuan Pei, Daochang Liu, Qian Luo, and Chang Xu (full)
	DM708: Inconsistency Distillation for Consistency: Enhancing Multi-View Clustering via Mutual Contrastive Teacher-Student Leaning	Dunqiang Liu, Xin Liu, Shu-Juan Peng, Lei Zhu, Zhen Cui, and Taihao Li (full)
	DM278: A Cognitive Solver with Autonomously Knowledge Learning for Reasoning Mathematical Answers	Jiayu Liu, Zhenya Huang, Xin Lin, Qi Liu, Jianhui Ma, and Enhong Chen (full)
	DM575: Hypernode: Entity Fusion for Data Tracebility and Link Prediction	Bingbing Dong, Jiao Li, Yi Zhu, Zan Zhang, Chenyang Bu, and Xindong Wu (full)
	DM714: Few-shot Partial Multi-label Learning with Data Augmentation	Yifan Sun, Yunfeng Zhao, Guoxian Yu, Zhongmin Yan, and Carlotta Domeniconi (full)
C3	C3: Recommender Ststems (5 full) (Remote)	
	DM628: AutoAssign: Automatic Shared Embedding Assignment in Streaming Recommendation	Fengyi Song, Bo Chen, Xiangyu Zhao, Huifeng Guo, and Ruiming Tang (full)
	DM239: Heterogeneous Graph Neural Network for Privacy-Preserving Recommendation	Yuecen Wei, Xingcheng Fu, Qingyun Sun, Hao Peng, Jia Wu, Jinyan Wang, and Xianxian Li (full)
	DM267: A Bi-directional Recommender System for Online Recruitment	Zhe-Rui Yang, Zhen-Yu He, Chang-Dong Wang, Pei-Yuan Lai, De-Zhang Liao, and Zhong-Zheng Wang (full)
	DM929: Which Companies are Likely to Invest: Knowledge-graph-based Recommendation for Investment Promotion	Chenyang Bu, jiawei zhang, Xingchen Yu, Le Wu, and Xindong Wu (full)
	DM483: One Person, One Model—Learning Compound Router for Sequential Recommendation	Zhiding Liu, Mingyue Cheng, Zhi Li, Qi Liu, and Enhong Chen (full)
C4	C4: Recommendation and Behavior Modeling (4 full + 2 short) (Remote)	
	DM306: Kernel-based Substructure Exploration for Next POI Recommendation	Wei Ju, Yifang Qin, Ziyue Qiao, Xiao Luo, Yifan Wang, Yanjie Fu, and Ming Zhang (full)
	DM1080: AutoAttention: Automatic Field Pair Selection for Attention in User Behavior Modeling	Zuowu Zheng, Xiaofeng Gao, Junwei Pan, Qi Luo, Guihai Chen, Dapeng Liu, and Jie Jiang (full)
	DM976: MixDec Sampling: A Soft Link-based Sampling Method of Graph Neural Network for Recommendation	Xiangjin Xie, Yuxin Chen, Ruipeng Wang, Xianli Zhang, Shilei Cao, Kai Ouyang, Zihan zhang, Hai-Tao Zheng, Buyue Qian, Bo Hu, Chengxiang Zhuo, and Zang Li (full)
	DM535: A Knowledge-Enhanced Framework for Imitative Transportation Trajectory Generation	Qingyan ZHU, Yize CHEN, Hao WANG, Zhenyu ZENG, and Hao LIU (full)
	DM512: Exploiting Hierarchical Correlations for Cross-City Cross-Mode Traffic Flow Prediction	Yan Chen, Jingjing Gu, Fuzhen Zhuang, Xinjiang Lu, and Ming Sun (short)
	DM592: Predicting Protein-Ligand Binding Affinity via Joint Global-Local Interaction Modeling	yang zhang, Gengmo zhou, Zhewei Wei, Hongteng Xu, and JiRong wen (short)

C5	C5: Graph Learning and Applications (4 full + 2 short) (Remote)	
	DM345: GRANDE: a neural model over directed multigraphs with application to anti-money laundering	Ruofan Wu, Boqun Ma, Hong Jin, Wenlong Zhao, Weiqiang Wang, and Tianyi Zhang (full)
	DM591: Neural Network Driven by Space-time Partial Differential Equation for Predicting Sea Surface Temperature	Taikang Yuan, Junxing Zhu, Kaijun Ren, Wuxin Wang, Xiang Wang, and Xiaoyong Li (full)
	DM789: Learning from the Multi-Level Abstraction of the Control Flow Graph via Alternating Propagation for Bug Localization	Yi-Fan Ma and Ming Li (full)
	DM527: Detach and Enhance: Learning Disentangled Cross-modal Latent Representation for Efficient Face-Voice Association and Matching	Zhenning Yu, Xin Liu, Yiu-ming Cheung, Minghang Zhu, Xing Xu, Nannan Wang, and Taihao Li (full)
	DM382: Contrastive Individual Treatment Effects Estimation	Xinshu Li and Lina Yao (short)
	DM250: MentalNet: Heterogeneous Graph Representation for Early Depression Detection	Ivan Mihov, Haiquan Chen, Xiao Qin, Wei-Shinn Ku, Da Yan, and Yuhong Liu (short)
C6	C6: Data and Model Aggregation (4 full + 2 short) (Remote)	
	DM925: Countering Modal Redundancy and Heterogeneity: A Self-Correcting Multimodal Fusion	Pengkun Wang, Xu Wang, Binwu Wang, Yudong Zhang, Lei Bai, and Yang Wang (full)
	DM771: TD3 with Reverse KL Regularizer for Offline Reinforcement Learning from Mixed Datasets	Yuanying Cai, Chuheng Zhang, Li Zhao, Wei Shen, Xuyun Zhang, Lei Song, Jiang Bian, Tao Qin, and Tieyan Liu (full)
	DM391: FedSkip: Combatting Statistical Heterogeneity with Federated Skip Aggregation	Ziqing Fan, Yanfeng Wang, Jiangchao Yao, Lingjuan Lyu, Ya Zhang, and Tian Qi (full)
	DM623: Contractible Regularization for Federated Learning on Non-IID Data	Zifan Chen, Zhe Wu, Xian Wu, Li Zhang, Jie Zhao, Yangtian Yan, and Yefeng Zheng (full)
	DM400: A Centralized-Distributed Transfer Model for Cross-Domain Recommendation Based on Multi-Source Heterogeneous Transfer Learning	Ke Xu, Ziliang Wang, Wei Zheng, Yuhao Ma, Chenglin Wang, Nengxue Jiang, and Cai Cao (short)
	DM415: Non-local Patch Mixup for Unsupervised Domain Adaptation	Jiujun He, Bin Liu, and Xuan Yang (short)
C7	S25: Recommender Systems and Graphs (3 full + 3 short) (Remote)	
	DM416: A Dynamic Variational Framework for Open-World Node Classification in Structured Sequences	Qin Zhang, Qincai Li, Xiaojun Chen, Peng Zhang, Shirui Pan, Philippe Fournier-Viger, and Joshua Zhexue Huang (full)
	DM709: Structural Robust Label Propagation on Homogeneous Graphs	Qiuting He, Jinsong Chen, Hao Xu, and Kun He (full)
	DM765: Multi-Graph Convolutional Recurrent Network for Fine-Grained Lane-Level Traffic Flow Imputation	Jingci Ming, Le Zhang, Wei Fan, Weijia Zhang, Yu Mei, Weicen Ling, and Hui Xiong (full)
	DM236: Empowerment-driven Policy Gradient Learning with Counterfactual Augmentation in Recommender Systems	Xiaocong Chen, Lina Yao, Xiaojun Chang, Siyu Wang, and Yong Li (short)
	DM1007: Decomposing Complementary and Substitutable Relations for Intercorporate Investment Recommendation	Le Dai, Yu Yin, Chuan Qin, Enhong Chen, and Hui Xiong (short)
	DM424: IGNiteR: News Recommendation in Microblogging Applications	Yuting Feng and Bogdan Cautis (short)
C8	S26: Graph Mining and Embedding (2 full + 4 short) (Remote)	

	DM406: Efficient Semi-Supervised Adversarial Training without Guessing Labels	Huimin Wu, William de Vazelhes, and Bin Gu (full)
	DM453: Towards Feature Distribution Alignment and Diversity Enhancement for Data-Free Quantization	Yangcheng Gao, Zhang Zhao, Richang Hong, Haijun Zhang, Jicong Fan, and Shuicheng Yan (full)
	DM658: Adversarial Inter-Group Link Injection Degrades the Fairness of Graph Neural Networks	Hussain Hussain, Meng Cao, Sandipan Sikdar, Denis Helic, Elisabeth Lex, Markus Strohmaier, and Roman Kern (short)
	DM986: Conflict Aware Pseudo Labeling via Optimal Transport for Entity Alignment	Qijie Ding, Daokun Zhang, and Jie Yin (short)
	DM387: Structure-Preserving Graph Representation Learning	Ruiyi Fang, liangjian Wen, Zhao Kang, and Jianzhuang Liu (short)
	DM488: Invariant Factor Graph Neural Networks	Zheng Fang, Ziyun Zhang, Guojie Song, Yingxue Zhang, Dong Li, Jianye Hao, and Xi Wang (short)
C9	C9: Graph and Relational Models (2 full + 4 short) (Remote)	
	DM473: Temporal Knowledge Graph Reasoning via Time-Distributed Representation Learning	Kangzheng Liu, Feng Zhao, Guandong Xu, Xianzhi Wang, and Hai Jin (full)
	DM498: Task-level Relations Modelling for Graph Meta-learning	Yuchen Zhou, Yanmin Shang, Yanan Cao, Chuan Zhou, Chuancheng Song, Fengzhao Shi, and Qian Li (full)
	DM755: Origin-Destination Traffic Prediction based on Hybrid Spatio-Temporal Network	Tingyang Chen, Lugang Nie, Jiwei Pan, Lai Tu, Bolong Zheng, and Xiang Bai (short)
	DM791: MRM2: Multi-Relationship Modeling Module for Multivariate Time Series Classification	Pengxiang Shi, Xuan Dang, Wenwen Ye, Zhipeng Li, and Zheng Qin (short)
	DM899: Robustness May be at Odds with Stability in Adversarial Training based Feature Selection?	Yue Liu and Yun Li (short)
	DM932: Higher-Order Masked Graph Neural Networks for Traffic Flow Prediction	Kaixin Yuan, Jing Liu, and Jian Lou (short)

Tutorials

Tutorial 1 (T1): Multi-Objective Optimization and Recommendations
https://moorecsys.github.io/ICDM2022/

Recommender systems have been widely applied to several domains and applications. Traditional recommender systems usually deal with a single objective, such as minimizing the prediction errors or maximizing the ranking of the recommendation list. There is an emerging demand for multi-objective optimization so that the development of recommendation models can take multiple objectives into consideration.

Currently, the multi-objective optimization methodologies have been well developed, and they have been reused in the area of recommender systems. This tutorial aims to provide a comprehensive introduction of the multi-objective optimization and the multi-objective recommender systems. More specifically, we identify the circumstances in which a multi-objective recommender system could be useful, summarize the methodologies and evaluation approaches in these systems, point out existing problems by a critical analysis, and provide guidelines for the usage of multi-objective optimization in recommender systems.

Tutorial 2 (T2): Fairness in Graph Mining: Metrics, Algorithms, and Applications

https://yushundong.github.io/ICDM_2022_tutorial.html

Graph data is ubiquitous in diverse real-world applications. To gain a deeper understanding of these graphs, graph mining algorithms have been playing a significant role over the years. Nevertheless, most graph mining algorithms lack fairness consideration. As a consequence, they could render discriminatory results against certain demographic subgroups or individuals. Such potential discrimination leads to an increasing societal concern about alleviating the bias exhibited in graph mining algorithms. In this tutorial, we provide a comprehensive overview of the recent research progress in measuring and mitigating the bias exhibited in graph mining algorithms. Specifically, we first introduce several widely-used fairness notions and the corresponding metrics. Then, we present an organized summary of existing techniques to debias graph mining algorithms. Furthermore, we demonstrate how different real-world applications benefit from these graph mining algorithms after debiasing. Finally, we provide insights on current research challenges and open questions to encourage further advances.

Tutorial 3 (T3): Mining of Real-world Hypergraphs: Patterns, Tools, and Generators

https://sites.google.com/view/hypergraph-tutorial-icdm

Group interactions are prevalent in various complex systems (e.g., collaborations of researchers and group discussions on online Q\&A sites), and they are commonly modeled as hypergraphs. Hyperedges, which compose a hypergraph, are non-empty subsets of any number of nodes, and thus each hyperedge naturally represents a group interaction among entities. The higher-order nature of hypergraphs brings about unique structural properties that have not been considered in ordinary pairwise graphs.

In this tutorial, we offer a comprehensive overview of a new research topic called *hypergraph mining*. Specifically, we first present recently revealed structural properties of real-world hypergraphs, including (a) static and dynamic patterns, (b) global and local patterns, and (c) connectivity and overlapping patterns. Together with the patterns, we describe advanced data mining tools used for their discovery. Lastly, we introduce simple yet realistic hypergraph generative models that provide an explanation of the structural properties.

Target audience: This tutorial is targeted at anyone interested in graph mining, social network analysis, or network science from researchers to practitioners from the industry.

Prerequisites: Basic knowledge of linear algebra and probability theory will be helpful. The tutorial will cover the necessary preliminaries and provide an intuitive overview of recent studies on the topic.

Tutorial 4 (T4): Recent and Emerging Trends in Tabular Data Generation

https://www.tulip.org.au/tutorials/ganblr/

Generative Adversarial network (GAN) model and its variants have shown to be effective in producing high-quality data in areas of Computer Vision, Text Mining and Natural Language Processing. GAN constitutes of two parts -- generator and discriminator, trained in an end-to-end manner in a game-theoretic manner. Tremendous success of GANs in producing high-quality structured data has inspired many researchers to utilize similar modelling for producing tabular data. Tabular data is a combination of apparently unrelated columns of types numeric, rank, and categorical features which makes the direct application of GAN-based deep learning methods quite challenging. This tutorial is aimed at discussing recent advancements in tabular data generation with GAN-style learning.

In this tutorial, we will start by providing a brief review of recent literature of various GAN-based techniques for tabular data generation. We will discuss various characteristics of tabular data and highlight the challenges of tabular data generation. We will also discuss the need for standard evaluation by proposing a centralized repository for comparing various tabular data generation methods. We will conclude this tutorial with a discussion of applications of tabular data

generation in privacy-preserving analytics, robustness analysis (concept drift analysis, adversarial attacks analysis) and anomaly detection.

Forums

Forum 1 (F1): Open Project Research Forum
https://www.cs.ucr.edu/~epapalex/icdm22-open-project-forum/

Open source projects are vital to the research community, since they enable reproducible research, providing confidence in existing published results and empowering further research by the community without the need to "reinvent the wheel". The data mining community has always been at the forefront of reproducible and open source research, which has enabled the rapid expansion of the field.

This year, ICDM is organizing the first open source project forum. The goal behind the forum is to highlight the importance of open projects and reproducible research, create a forum where researchers can showcase their open projects and discuss challenges, ideas, and future directions, and ultimately celebrate the transformative nature of open source projects in data mining.

The forum invites contributions from the community, as described in the Call for Papers below. In addition to contributed papers, the forum is planning to feature keynote talks by researchers in the community who have successfully created, released, and maintained prominent and high-impact open projects, who will discuss their experience and lessons-learned, followed by a panel discussion (pending speaker availability).

Forum 2 (F2): Women in Science Resesrach Forum