

Table of Contents

Welcome to PAKDD2022	1
General Chairs' Preface	2
Program Chairs' Preface	3
Steering Committee	4
Organizing Committee	5
Guidelines for Onsite Attendance	7
Guidelines for Online Attendance	8
Agenda	9
Keynote Speakers	13
Details on Tutorials & Workshop	16
Details on Sessions	25



Welcome to PAKDD2022

Welcome to the 26th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD2022), which will be held in Chengdu, China on May 16-19, 2022. The PAKDD is one of the longest established and leading international conferences in the areas of data mining and knowledge discovery. It provides an international forum for researchers and industry practitioners to share their new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, visualization, decision-making systems, and the emerging applications.

Organizer



Southwest Jiaotong University (SWJTU) was founded to train China's own railway engineers. Since our early days, we have been cultivating talents for the rejuvenation of Chinese nation and persevering in self-improvement with the motto "Diligence, Ambition, Tenacity and Loyalty". By training our graduates to master cutting-edge technologies and to strive for excellence, we have been shaping the development of China's rail transit industry from the beginning. We have more than 300,000 graduates whose ideas, innovations and actions have generated a positive impact in China and the world.

As an outstanding academic community, it has always been our mission to serve the country and the world through scientific education, research, and innovation, to solve fundamental scientific problems and to meet the challenges of today and tomorrow. SWJTU is situated in Chengdu, Sichuan's capital, and is known as the home of the beloved Panda Bear. Our hometown, a "garden city" is regarded as one of the metropolitan cities with the highest quality of life in China. Furthermore, since it is a key node city along the "Belt and Road", SWJTU is well positioned to adopt a global vision, with partners, cooperation projects, and alumni spanning the globe. Our growing scientific community attracts talents from all over the world and our students have opportunities to study, research, and work with outstanding and multidisciplinary teachers, alumni, and partners. Both rigorous and dynamic, the academic environment provides outstanding opportunities for the integration of theory and practice. As a member of the innovative, cooperative, and multidisciplinary community at SWJTU, you will have the chance to work in a community that serves as a key node of the Belt and Road Initiative, address the greatest global engineering challenges of today, and make a positive impact on the field of higher education with us. In addition, you can experience and enjoy our welcoming culture, highly praised by our international guests and visitors.



General Chairs' Preface

On behalf of the Organizing Committee, it is our great pleasure to welcome you to the 26th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD2022), held in Chengdu, China, during May 16-19, 2022. Starting in 1997, PAKDD has long established itself as one of the leading international conferences in data mining and knowledge discovery. PAKDD provides an international forum for researchers and industry practitioners to share their new ideas, original research results, and practical development experiences from all Knowledge Discovery and Data Mining (KDD) related areas. In response to the COVID-19 pandemic and the need for social distancing, PAKDD2022 was held as a hybrid conference for both online and onsite attendees.

Our gratitude goes first and foremost to the researchers, who submitted their work to PAKDD2022. We would like to deliver our sincere thanks for their efforts in research, as well as in preparing high-quality presentations. We also thank all the collaborators and sponsors for their trust and cooperation. It is our great honor that three eminent keynote speakers joined the conference: Jian Pei (Simon Fraser University, Canada), Bernhard Schölkopf (Max Planck Institute for Intelligent Systems, Germany) and Ji-Rong Wen (Renmin University, China). They were extremely professional and have high reputations in their respective areas. We enjoyed their participation and talks, which made the conference one of the best academic platforms for knowledge discovery and data mining.

We would like to express our sincere gratitude to the contributions of Steering Committee members, Organizing Committee members, Program Committee members and anonymous reviewers, led by Program Committee Co-chairs: Joao Gama (University of Porto), Tianrui Li (Southwest Jiaotong University), and Yang Yu (Nanjing University).

We are also grateful for the hosting organization Southwest Jiaotong University which is continuously providing institutional and financial support to PAKDD2022. We feel beholden to the PAKDD Steering Committees for their constant guidance and sponsorship of manuscripts.

Finally, our sincere thanks go to all the participants and volunteers. We hope all of you enjoyed PAKDD2022.

Enhong Chen Yu Zheng



Program Chairs' Preface

It is our great pleasure to present at the 26th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD2022) as the Program Committee Chairs. PAKDD is one of the longest established and leading international conferences in the areas of data mining and knowledge discovery. It provides an international forum for researchers and industry practitioners to share their new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, big data technologies and foundations.

This year PAKDD received 627 submissions, among which 69 submissions were rejected at a preliminarily stage due to the policy violations. There were 320 Program Committee members and 45 Senior Program Committees members involved in reviewing process. Each submission was reviewed by at least three different reviewers. Over 67% of those submissions were reviewed by four or more reviewers. Eventually, 121 submissions were accepted and recommended to be published, resulting in an acceptance rate of 19.30%. Out of these, 29 submissions were about applications, 4 submissions were related to big data technologies, 46 submissions were on data science and 42 submissions were about foundations. We would like to appreciate all PC members and reviewers, who offered a high-quality program with diligence on PAKDD2022.

The conference program featured keynote speeches from distinguished researchers in the community, most influential paper talks, cutting-edge workshops and comprehensive tutorials.

We wish to sincerely thank all PC members and reviewers for their invaluable efforts in ensuring a timely, fair, and highly effective PAKDD2022 program.

Joao Gama Tianrui Li Yang Yu



Steering Committee

- Longbing Cao, University of Technology Sydney
- Ming-Syan Chen, Nanyang Technological University
- David Cheung, University of Hong Kong
- Gill Dobbie, University of Auckland
- Joao Gama, University of Porto
- Zhiguo Gong, University of Macau
- Tu Bao Ho, Japan Advanced Institute of Science and Technology
- Joshua Z. Huang, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
- Masaru Kitsuregawa, Tokyo University
- Rao Kotagiri, University of Melbourne
- Jae-Gil Lee, Korea Advanced Institute of Science & Technology
- **Ee-Peng Lim**, Singapore Management University
- Huan Liu, Arizona State University
- Hiroshi Motoda, AFOSR/AOARD and Osaka University
- Jian Pei, Simon Fraser University
- Dinh Phung, Monash University
- P. Krishna Reddy, International Institute of Information Technology, Hyderabad (IIIT-H)
- Kyuseok Shim, Seoul National University
- Jaideep Srivastava, University of Minnesota
- Thanaruk Theeramunkong, Thammasat University
- Vincent S. Tseng, NCTU
- Takashi Washio, Osaka University
- Geoff Webb, Monash University
- Kyu-Young Whang, Korea Advanced Institute of Science & Technology
- Graham Williams, Australian National University
- Min-Ling Zhang, Southeast University
- Chengqi Zhang, University of Technology Sydney
- Ning Zhong, Maebashi Institute of Technology
- Zhi-Hua Zhou, Nanjing University



Organizing Committee

Honorary Co-Chairs

- Dan Yang, Southwest Jiaotong University
- Zhi-Hua Zhou, Nanjing University

General Co-Chairs

- Enhong Chen, University of Science and Technology of China
- Yu Zheng, JD.com

Program Committee Co-Chairs

- Joao Gama, University of Porto
- Tianrui Li, Southwest Jiaotong University
- Yang Yu, Nanjing University

Workshop Co-Chairs

- Gill Dobbie, University of Auckland
- Can Wang, Griffith University

Tutorial Co-Chairs

- Gang Li, Deakin University
- Tanmoy Chakraborty, Indraprastha Institute of Information Technology Delhi

Local Arrangement Co-Chairs

- Yan Yang, Southwest Jiaotong University
- Chuan Luo, Sichuan University
- Xin Yang, Southwestern University of Finance and Economics



Organizing Committee

Sponsor Co-Chair

Xiaobo Zhang, Southwest Jiaotong University

Publicity Co-Chairs

- Xiangnan Ren, Group 42
- Hao Wang, Zhejiang Lab
- Junbo Zhang, JD.com
- Chongshou Li, Southwest Jiaotong University

Proceedings Chair

• Fei Teng, Southwest Jiaotong University

Web and Content Co-chairs

- Xiaole Zhao, Southwest Jiaotong University
- Zhen Jia, Southwest Jiaotong University

Registration Chairs

- Hongmei Chen, Southwest Jiaotong University
- Jie Hu, Southwest Jiaotong University
- Yanyong Huang, Southwestern University of Finance and Economics



Guidelines

FOR ONSITE ATTENDANCE

Conference Venue

Crowne Plaza Chengdu West | 成都新希望高新皇冠假日酒店
 Address: No. 1 Xixin Avenue, Pidu District, Chengdu City, Sichuan Province, China | 四川省成都市高新西区西芯大道 1 号

For Oral Presentation

- The duration of a presentation slot is 17 minutes. Please target your lecture for a duration of about 15 minutes for the presentation plus about 2 minutes for questions from the audience.
- Your punctual arrival and active involvement in each session will be highly appreciated.
- Get your presentation PPT or PDF files prepared and backed up.
- Laptop (with MS-Office & Adobe Reader), projector & screen, laser sticks will be provided by the conference organizer.

Duplication/ Recording

 Unauthorized photography, audio taping, video recording, digital taping or any other form of duplication is prohibited in the conference.

Security

 Please ensure that you take all items of value with you at all times when leaving a room. Do not leave bags or laptops unattended.

Name Badge

• For security purposes, delegates, speakers, exhibitors and staff are required to wear their name badge to all sessions and social functions. Entrance into sessions is restricted to registered delegates only. If you misplace your name badge, please see the staff at the registration desk to arrange a replacement.

Safety & Well-being

- You are required to wear face masks in public areas.
- Good hand hygiene is one of the most important steps individuals can take to protect themselves.
- Attendees are suggested to have a body temperature check prior to arriving at the venue.
- Signs of COVID-19 and/or with a fever of 37.3° or greater should be reported immediately.
- You are requested to present your health code before entering the conference venue.
- Scan the QR code (on the right) to obtain your health code.





Guidelines

FOR VIRTUAL TALKS

Platform: Zoom

- For Users from mainland China please download: www.zoom.com.cn/download
- For General Users please download: https://zoom.us/support/download
- Zoom Help Center: https://support.zoom.us

Time Zone

- China Standard Time (CST) UTC/GMT+08:00
- Please make sure that both the clock and the time zone on your computer are set to the correct China
 Time

Device

- A computer with an internet connection (wired connection recommended)
- USB plug-in headset with a microphone (recommended for optimal audio quality)
- Webcam (optional): built-in or USB plug-in

Environment

- Quiet Environment
- Stable Internet Connection
- Proper lighting

Sign In and Join

- Join a meeting without signing in: A Zoom account is not required if you join a meeting as a participant,
 but you cannot change the virtual background or edit the profile picture
- Sign in with a Zoom account: All the functions are available

Voice Control Rules

- The host will mute all participants while entering the meeting.
- Speakers can unmute microphone when it is his or her turn for presentation.

Conference Recording

- The whole conference will be recorded. We appreciate your proper behavior and appearance.
- The recording will be used for the conference reports among the committee. It won't be distributed to or shared with anyone else, and it shall not be used for commercial or illegal purpose. It will only be recorded by the staff; the presenters are not allowed to record.



May 16, 2022 (Monday)

AGENDA | ONSITE

Time	Event	
	Crowne Plaza Chengdu West	
10:00-17:00	Sign-In Onsite to Collect Materials	Lobby / 1F
18:00-20:00	Reception	Restaurant / 1F

AGENDA | ONLINE

Time

Zoom ID:	815 9034 3067 (Room 1)	829 6485 3183 (Room 2)	822 3121 1337 (Room 3)
09:00-12:00	Tutorial 1 Unrevealing Data Correlations with Self-Supervised Learning	Tutorial 2 Recent Advances in Causality-based Recommendation	
14:00-17:00	Tutorial 3 Harnessing the Power of GAN for Tabular Data Generation	Tutorial 4 Online clustering: algorithms, evaluation, metrics, applications and benchmarking using River	Workshop 2nd International Workshop on Smart & Precise Agriculture (WSPA 2.0)

Zoom ID:	839 4588 3358 (Room 5)	
10:00-12:00	Online Test All the keynote speakers/presenters/session chairs are suggested to attend the test session;	
14:00-17:00	you can join the room at your appropriate local time between 10:00-12:00 & 14:00-17:00 (CST), May 16, 2022.	



May 17, 2022 (Tuesday)

AGENDA | ONSITE

Time	Event	
	Diamond Ballroom / 2F 璀璨宴会厅 Online Room: 815 9034 3067	
08:30-09:00	Opening Ceremony Opening remarks: Dan Yang Opening remarks from general chair: Zheng Yu Opening remarks from program committee chair: Yang Yu	Host: Tianrui Li
09:00-10:00	Keynote Speech I: Jian Pei, Simon Fraser University, Canada	Host: Yu Zheng
10:00-10:30	Group Photo & Coffee Break	
10:30-12:30	Session: T1-Application	Chair: Chuan Luo
12:30-13:30	Lunch @ Restaurant / 1F	
13:30-15:50	Session: T2-Data Science	Chair: Xin Yang
15:50-16:10	Coffee Break	
16:10-18:30	Session: T3-Foundation	Chair: Fan Min
19:00-20:30	Dinner @ Restaurant / 1F	

AGENDA | ONLINE

Time	Event	
Zoom ID:	829 6485 3183 (Room 2)	822 3121 1337 (Room 3)
10:30-12:30	Session: V1-Data Science	Session: V2-Foundation
12:30-13:30	Lunch Break	
13:30-15:50	Session: V3-Application	Session: V4-Foundation
16:10-18:30	Session: V5-Data Science	Session: V6-Application



May 18, 2022 (Wednesday)

AGENDA | ONSITE

Time	Event	
	Diamond Ballroom / 2F 璀璨宴会厅 Online Room 1: 815 9034 3067	
08:30-09:30	Keynote Speech II: Ji-Rong Wen, Renmin University of China, China	Host: Enhong Chen
09:30-10:40	Session: T4-Data Science	Chair: Zhi Xu
10:40-11:00	Coffee Break	
11:00-12:30	Session: T5-Foundation	Chair: Chongshou Li
12:30-13:30	Lunch @ Restaurant / 1F	
13:30-17:30	Academic Visit	
19:00-20:30	Banquet @ Restaurant / 1F	

AGENDA | ONLINE

Time	Event	
Zoom ID:	829 6485 3183 (Room 2)	822 3121 1337 (Room 3)
09:30-10:40	Session: V7-Data Science	Session: V8-Foundation
10:40-11:00	Break	
11:00-12:30	Session: V9-Data Science	Session: V10-Foundation



May 19, 2022 (Thursday)

AGENDA | ONLINE

Time	Event		
Zoom ID:	815 9034 3067 (Room 1)	829 6485 3183 (Room 2)	822 3121 1337 (Room 3)
09:00-10:30	Session: V11-Data Science	Session: V12-Data Science	Session: V13-Foundation
10:30-10:50	Break		
10:50-12:00	Session: V14-Data Science	Session: V15-Application	Session: V16-Application
09:30-12:00	Committee Meeting	Committee Meeting	
12:00-14:00	Lunch Break	Lunch Break	
Zoom ID:	815 9034 3067 (Room 1)		
14:00-15:00	Keynote Speech III: Bernhard Schölkopf, Max Planck Institute for Intelligent Systems, Germany Host: Tianrui Li		Host: Tianrui Li
15:00-15:20	Break		
15:20-16:00	PAKDD Most Influential Paper Talk Host: Yang Yu		
16:00-16:30	Closing Ceremony Host: Yang Yu		



Keynote Speaker



Prof. Jian PeiSimon Fraser University, Canada

Jian Pei is a Professor at Simon Fraser University. His research focuses on data science, big data, data mining, database systems, and information retrieval. His expertise is in developing effective and efficient data analysis techniques for novel

data intensive applications and transferring his research results to industry products and business practice. He is recognized as a Fellow of the Royal Society of Canada (Canada's national academy), the Canadian Academy of Engineering, ACM, and IEEE. Since 2000, he has published one textbook, two monographs and over 300 research papers in refereed journals and conferences, which have been cited extensively. He was the editor-in-chief of the IEEE Transactions of Knowledge and Data Engineering (TKDE) in 2013-16, the chair of ACM SIGKDD in 2017-2021, and the organizers of many conferences. He received a few prestigious awards, including the 2017 ACM SIGKDD Innovation Award, the 2015 ACM SIGKDD Service Award, the 2014 IEEE ICDM Research Contributions Award, the British Columbia Innovation Council 2005 Young Innovator Award, an IBM Faculty Award, a KDD Best Application Paper Award, and an ICDE Influential Paper Award.

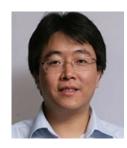
SPEECH TOPIC

Exact, Concise, and Consistent Data Driven Interpretation

Interpretability and explainability are at the core in human being's pursuit of new knowledge. At the same time, interpretation in data analytics and data mining is challenging in many ways, such as the complexity of models to be interpreted, the difficulty in knowledge elicitation, the expectation of embodying interpretation, and the need of many kinds of knowledge. In this talk, I will present our systematic research on exact, concise, and consistent data driven interpretation for database and data mining tasks. I will illustrate our principles and techniques using several application examples, including multidimensional skyline queries (aka pareto optima) in databases, piece-wise linear neural networks in classification, and KS-tests in statistics. I will also discuss the promises and challenges of data driven interpretation for future work.



Keynote Speaker



Prof. Ji-Rong WenRenmin University of China, China

Ji-Rong Wen is a full professor, the dean of School of Information, and the executive dean of Gaoling School of Artificial Intelligence at Renmin University of China (RUC). He has been working in the big data and Al areas for many years. He

is the PC Chair of SIGIR 2020 and the Associate Editor of ACM TOIS and IEEE TKDE. He worked at Microsoft Research Asia (MSRA) for 14 years and once was a senior researcher and the group manager of the Web Search and Mining Group. He was elected as a National Distinguished Professor in 2013 and Beijing's Distinguished Young Scientist in 2018. He is a Chief Scientist of Beijing Academy of Artificial Intelligence.

SPEECH TOPIC

Wenlan - A Large-scale Multi-modal Pre-trained Model

Abstract: In this talk, I will introduce our recent work on a large-scale multi-modal pre-trained model named Wenlan. Wenlan was trained in a self-supervised way on huge datasets containing billions of image-text pairs collected from the internet and millions of short videos. We have applied Wenlan to a number of downstream tasks and demonstrated its superiority and versatility. Moreover, I will share some insights by further exploring and exploiting the Wenlan model, which verify that multi-modal pre-training is a promising way to get better representation, structure and knowledge like human beings.



Keynote Speaker



Prof. Bernhard SchölkopfMax Planck Institute for Intelligent Systems, Germany

Bernhard Schölkopf's scientific interests are in machine learning and causal inference. He has applied his methods to a number of different fields, ranging from biomedical problems to computational photography and astronomy. Bernhard has

researched at AT&T Bell Labs, at GMD FIRST, Berlin, and at Microsoft Research Cambridge, UK, before becoming a Max Planck director in 2001. He is a member of the German Academy of Sciences (Leopoldina), has (co-)received the Royal Society Milner Award, the Leibniz Award, the Koerber European Science Prize, and the BBVA Foundation Frontiers of Knowledge Award, He is Fellow of the ACM and of the CIFAR Program "Learning in Machines and Brains", an Amazon Distinguished Scholar, and holds a Professorship at ETH Zurich. Bernhard co-founded the series of Machine Learning Summer Schools, and helped build the Journal of Machine Learning Research, an early development in open access and today the field's flagship journal.

SPEECH TOPIC

From Statistical to Causal Machine Learning

In machine learning, we use data to automatically find dependences in the world, with the goal of predicting future observations. Its methods build on statistical dependences, but one can try to go beyond this, assaying underlying causal structures.

Causal models may be more robust to changes that occur in real world datasets and thus play a central role in addressing some of the hard open problems of the field. We discuss implications of causal models for machine learning, as well as connections in the opposite direction, including the prospects of causal representation learning.



Tutorial 1

Unrevealing Data Correlations with Self-Supervised Learning

Sonali Agarwal, Sanjay Kumar Sonbhadra, Narinder Singh Punn

Time: 09:00-12:00

Zoom: 815 9034 3067 (Room 1)

Abstract: Deep learning has brought the most profound contribution towards living standards by addressing several real-world problems. The continuous advancements in the approaches for classification, localization, segmentation, detection, etc. have extended its application spectrum across various domains. To accomplish such tasks, the models are required to be trained using a huge amount of annotated or labeled data. However, the generation of the annotations for such huge data requires expert analysts and extensive manual efforts. It is a tedious and expensive task, while also being vulnerable to error. Self-supervised learning is an emerging technology that advances to address this issue by effectively closing the gap with fully supervised methods. Here, the aim is to perform pre-training with an unsupervised strategy for learning useful and better representations of the data samples. The pre-trained model is then fine-tuned with limited annotated samples to adopt the actual task such as segmentation, classification, etc. In light of this, the present tutorial discovers various self-supervised learning strategies that could be utilized to improve the performance of the deep learning models.

SPEAKER BIO

Dr. Sonali Agarwal is working as an Associate Professor in the Information Technology Department of Indian Institute of Information Technology (IIIT), Allahabad, India. She received her Ph. D. Degree at IIIT Allahabad and joined as faculty at IIIT Allahabad, where she has been teaching since October 2009. She holds Bachelor of Engineering (B.E.) degree in Electrical Engineering from Bhilai Institute of Technology, Bhilai, (C.G.) India and Masters of Engineering (M.E.) degree in Computer Science from Motilal Nehru National Institute of Technology (MNNIT), Allahabad, India Her main research interests are in the areas of Artificial Intelligence and Big Data. She is the head of Big Data Analytics Lab at IIIT Allahabad, India.

Dr. Sanjay Kumar Sonbhadra is presently working as Assistant Professor in the Computer Science and Engineering Department of ITER, Shiksha 'O' Anusandhan, Bhubaneswar, Odisha, India. He is mainly working on One-class classification, Anomaly detection, Target class guided dimensionality reduction and training sample selection techniques and Big data analytics. During 2017-2021, he worked as a senior member of "Big Data Analytics Lab" at IIIT Allahabad, India. He has published many articles in the area of machine learning applications to address recent challenges of COVID-19.



Mr. Narinder Singh Punn is working as a Teaching Research Assistant (TRA) in the Information Technology Department of Indian Institute of Information Technology (IIIT), Allahabad, India. Narinder's main research includes Medical Imaging segmentation, Deep learning and Artificial Intelligence techniques in healthcare. He is a senior member of "Big Data Analytics Lab" at IIIT Allahabad, India. His recent publications cover applications of deep learning in the detection and prevention of COVID-19, while also exploiting the potential of self-supervised learning in biomedical image segmentation.



Tutorial 2

Recent Advances in Causality-based Recommendation

Qian Li, Gang Li, Guangdong Xu

Time: 09:00-12:00

Zoom: 829 6485 3183 (Room 2)

Abstract: Recommender system (RS) has become a panacea to assist users in discovering preferred contents from the massive information provided. With the aim of delivering high-quality personalized services and items to users, RS has received extensive attention from both research communities and industries over the last few years. Prior efforts extensively rely on complex machine learning or latent representation models to build the highly effective and accurate recommender system. Recently, providing robust and understandable explanations to fully address the problem of why an item is recommended has gained a lot of momentum in both industrial and research communities. However, this objective faces two significant challenges based on the ubiquitous user item interaction data.

- 1. Bias problem: user-item interaction usually exhibits bias that is entangled with users' real interests, ultimately degrading the recommendation performance. Bias is common in RS including exposure bias, popularity bias, selection bias and so on. Take selection bias for example, in movie recommendation, users are more likely to watch movies that are watched by many people, which however is due to users' conformity to other people, rather than stemming from users' real interests. Therefore, it is essential to capture users' pure interests that are independent of the bias and thus can be leveraged to build high-quality recommender models.
- 2. Non-causal explanations: existing methods explore explanations by correlations and not causality, which leads to inconsistency of explanations. Since users can behave differently depending on the recommended policies, it requires the recommendation methods to be robust and highly adaptive to investigate the causes for user's true interests to the ever-changing user and recommender environments.

Causality learning has become a new promising direction to address these two challenges. In a broader sense, causality based recommendation attempts to generate both high-quality recommendations and intuitive explanations for users, either in a post-hoc way or by constructing an inherent interpretable model. In this tutorial, we will start with two challenges faced by recommendation including bias issue and non-causal interpretability. Then, we will introduce the basic concept of causality learning and how it is used for recommendations. Particularly, various causality-based approaches to address the challenges of bias and non-causal interpretability will be introduced. Finally, we will figure out the existing open problems in causality-based recommendation and point out the future work.



SPEAKER BIO

Qian Li is a Lecturer (a.k.a., Assistant Professor) at the School of Electrical Engineering, Computing and Mathematical Sciences (EECMS), Curtin University, Australia. Before that, she has been a Postdoc Research Fellow in the Data Science and Machine Intelligence (DSMI) Lab at University of Technology Sydney (UTS) since July 2019. She received her Ph.D. degree in Computer Science from the Institute of Information Engineering, Chinese Academy of Science (CAS). Before that, she got a Master degree (Master by Research) in Computer Science at Shandong University. She also won the Luxembourg Scholarship and completed a 2nd M.S. degree (Master by Research) in Computer Science at the University of Luxembourg. Her recent interests lie in causal machine learning and exploring causal reasoning insights to tackle challenging problems in machine learning, such as robustness and interpretability. Besides, her research focuses on utilizing advanced mathematical tools (such as optimal transport and Riemannian geometry) to alleviate the challenges of computer vision tasks.



Tutorial 3

Harnessing the Power of Generative Adversarial Networks Style Learning for Tabular Data Generation

Nayyar A. Zaidi, Yishuo Zhang, and Gang Li

Time: 14:00-17:00 Zoom: 815 9034 3067 (Room 1)

Abstract: 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.

SPEAKER BIO

Nayyar A. Zaidi is currently a Senior Lecturer at Deakin University and Lead Research Scientist at DataScienceWorks Research. He received the B.S. degree in computer science and engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2005, and the Ph.D. degree in Artificial Intelligence from Monash University, Melbourne, VIC, Australia, in 2011. He worked as a Research Fellow, a Lecturer, and a Research Fellow, from 2011 to 2013, from 2013 to 2014, and from 2014 to 2017, respectively, at the Faculty of Information Technology, Monash University. From 2017 to 2019, he worked as Research Scientist at Credit AI (Trusting Social) Melbourne Lab. His research interests include effective feature engineering, explainable model, uncertainty prediction, and reinforcement learning. He is also interested in practical data science, machine learning engineering, and data science training. He was a recipient of the Gold Medal for graduating top of the class at the University of Engineering and Technology.



Yishuo Zhang received his B.S. degree in computer science from the University of Zhengzhou, China in 2010, and the M.S degree in information technology from Monash University, Melbourne, VIC, Australia, in 2013. He currently is the second year Ph.D. student at the School of Information Technology, Deakin University and his research interests include big data feature engineering, tabular data generation, the trust-able and explainable model and tourism demand forecasting.

Gang Li, IEEE senior member, received his Ph.D. in computer science in 2005. He joined the School of Information Technology at Deakin University (Australia) as an associate lecturer (2004-2006), lecturer (2007-2011), senior lecturer (2012-2016). His research interests are in the area of data mining, machine learning, and business intelligence. He serves on the IEEE Data Mining and Big Data Analytics Technical Committee (2017-2018 Vice Chair), and IEEE Enterprise Information Systems Technical Committee, IEEE Enterprise Architecture and Engineering Technical Committee, and serves as chair for IEEE Task force on Educational Data Mining (2020-2023 Chair). He acts as an associate editor for Decision Support Systems (Elsevier), IEEE Access (IEEE), Journal of Travel Research (Sage), and Information Discovery & Delivery (Emerald), and Human-Centric Computing and Information Sciences (Springer) etc. He has been the guest editor for IEEE Access, the Chinese Journal of Computer, Journal of Networks, Future Generation Computer Systems (Elsevier), Concurrency and Computation: Practise and Experience (Wiley) and Enterprise Information Systems (Taylor & Francis). He has co-authored 8 papers that won best paper prizes, including KSEM 2018 Best Paper award, IFITT Journal Paper of the Year (2017, 1st prize), IEEE Trustcom 2016 best student paper award, Journal Paper of the Year (2015, 3rd award) from IFITT, the PAKDD2014 best student paper award, ACM/IEEE ASONAM2012 best paper award, the 2007 Nightingale Prize by Springer journal Medical and Biological Engineering and Computing. He has also conducted research projects on tourism and hospitality management. He served on the Program Committee for over 150 international conferences in artificial intelligence, data mining, machine learning, tourism and hospitality management, and is a regular reviewer for International Journals in the areas of data science, privacy protection, recommendation system, and business intelligence.



Tutorial 4

Online Clustering: Algorithms, Evaluation, Metrics, Applications and Benchmarking using River

Jacob Montiel, Hoang-Anh Ngo, Minh-Huong Le-Nguyen and Albert Bifet

Time: 14:00-17:00

Zoom: 829 6485 3183 (Room 2)

Abstract: Online clustering algorithms play a critical role in data science, especially with the advantages regarding time, memory usage and complexity, while maintaining a high performance compared to traditional clustering methods. This tutorial serves, first, as a survey on online machine learning and, in particular, data stream clustering methods. This introduction will be then put into the context with River, a go-to Python library merged between Creme and scikit-multiflow. We propose applications and settings for benchmarking, using real-world problems and datasets.

SPEAKER BIO

Jacob Montiel is a research fellow at the University of Waikato in New Zealand and the core developer and maintainer of river. His research interests are in the field of machine learning for evolving data streams. Prior to focusing on research, Jacob led development work for onboard software for aircraft and engine's prognostics at GE Aviation; working in the development of GE's Brilliant Machines, part of the loT and GE's approach to Industrial Big Data.

Website: https://jacobmontiel.github.io/

Albert Bifet is a Professor of Al and the Director of the Te Ipu o te Mahara Al Institute at University of Waikato, and Professor of Big Data at Data, Intelligence and Graphs (DIG) LTCI, T'el'ecom Paris. Problems he investigate are motivated by large scale data, the Internet of Things (IoT), and Big Data Science. He co-leads the open source projects MOA (Massive On-line Analysis), Apache SAMOA (Scalable Advanced Massive Online Analysis) and StreamDM.

Website: http://albertbifet.com

Hoang-Anh Ngo is currently a Research Assistant at Data, Intelligence and Graph (DIG), LCTI, Telecom Paris He is also a core developer and maintainer of River, the machine learning library in Python for data streams. His research interests lies in the field of machine learning for evolving data stream, particularly in online clustering and classification algorithms. Previously, he joined the team of IT Specialists in COVID-19 task force, formed by the Ministry of Health of Vietnam as a Epidemiological Modelling Unit head.



Minh-Huong LE NGUYEN is a third-year doctoral student at Telecom Paris, Institut Polytechnique de Paris in France. Her doctoral research focuses on 2 the applications of machine learning on data streams to implement predictive maintenance in the railway industry. She received her Bachelor's degree in Computer Science at University Pierre and Marie Curie (France) in 2013, then she graduated from Telecom Paris (France) with a Master's degree in Data science in 2019.



2nd International Workshop on Smart & Precise Agriculture (WSPA 2.0)

Workshop website: https://creds.iitpkd.ac.in/wspa

Time: 14:00-17:00

Zoom: 822 3121 1337 (Room 3)

Objective of the workshop is to inform, encourage, and showcase the important and emerging area of smart agriculture which can have a significant impact in the economy and life of a huge population. More importantly, the workshop aims to draw the attention of researchers in the area where there are novel and hard problems in the intersection of machine learning, edge devices, sensor devices, internet technologies and agriculture.

Invited Talk

Prof. Sruti Das Choudhury

Department of Computer Science and Engineering, University of Nebraska Lincoln

Accepted Papers

- Paddy Doctor: A visual image dataset for paddy disease classification Petchiammal A, Briskline Kiruba S, Murugan D and Pandarasamy Arjunan
- Self-Supervised Anomaly Detection of Rogue Soil Moisture Sensors Boje Deforce, Bart Baesens, Jan Diels and Estefanía Serral Asensio
- Effect of Rainfall and Land Use Change Dynamics on Groundwater Fluctuation: A Case Study of Dehradun

Ravi Kant Bhardwaj, Shivangi Sharma and Deepak Kumar

Organizing Committee

Dr. Mrinal Das, IIT Palakkad

Dr. Ayan Chaudhury, IIT Kharagpur

Dr. Sahely Bhadra, IIT Palakkad

Dr. Satyajit Das, IIT Palakkad



Session | T1-Application

May 17/Tuesday

Room: Diamond Ballroom / 2F | 璀璨宴会厅

Zoom ID: 815 9034 3067

Time: 10:30-12:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Chuan Luo

#538	Title: Rethinking Adjacent Dependency in Session-based Recommendations Authors: Qian Zhang, Shoujin Wang, Wenpeng Lu, Chong Feng, Xueping Peng, Qingxiang Wang Presenter: Qian Zhang Affiliation: Qilu University of Technology (Shandong Academy of Sciences)
#277	Title: Causal Enhanced Uplift Model Authors: Xiaofeng He, Guoqiang Xu, Cunxiang Yin, Zhongyu Wei, Yuncong Li, Yancheng He, and Jing Cai Presenter: Xiaofeng He Affiliation: Fudan University
#365	Title: Learning Discriminative Representation base on Attention for Uplift Authors: Guoqiang Xu, Cunxiang Yin, Yuchen Zhang, Yuncong Li, Yancheng He, Jing Cai, Zhongyu Wei Presenter: Guoqiang Xu Affiliation: Tencent
#113	Title: Input Enhanced Logarithmic Factorization Network For CTR Prediction Authors: Xianzhuang Li, Zhen Wang, Xuesong Wu, Bo Yuan, Xueqian Wang Presenter: Xianzhuang Li Affiliation: Tsinghua University
#656	Title: A Novel Protein Interface Prediction Framework via Hybrid Attention Mechanism Authors: Haifang Wu, Shujie Luo, Weizhong Zhao, Xingpeng Jiang, Tingting He Presenter: Weizhong Zhao Affiliation: Central China Normal University
#139	Title: Bribery in Rating Systems: A Game-Theoretic Perspective Authors: Xin Zhou, Shigeo Matsubara, Yuan Liu, Qidong Liu Presenter: Qidong Liu Affiliation: Zhengzhou University
#499	Title: Improve Chinese Spelling Check by Reevaluation Authors: Shuai Wang, Lin Shang Presenter: Shuai Wang Affiliation: Nanjing University



Session | T2 – Data Science

May 17/Tuesday

Room: Diamond Ballroom / 2F | 璀璨宴会厅

Zoom ID: 815 9034 3067

Time: 13:30-15:50 (Duration for Each Presentation: 17 minutes)

Session Chair: Xin Yang

#445	Title: Overcoming Forgetting in Local Adaptation of Federated Learning Model Authors: Shunjian Liu, Xinxin Feng, Haifeng Zheng Presenter: Shunjian Liu Affiliation: Fuzhou University
#367	Title: Domain-level Pairwise Semantic Interaction for Aspect-Based Sentiment Classification Authors: Zhenxin Wu, Jiazheng Gong, Kecen Guo, Guanye Liang, Qingliang Chen, Bo Liu Presenter: Zhenxin Wu Affiliation: Jinan University
#168	Title: LCAN: Light Cross-Attention Network for Collaborative Filtering Recommendation Authors: Lin Liu, Wei Zhou, Junhao Wen, Yihao Zhang, Yu Wang, Hanwen Zhang Presenter: Lin Liu Affiliation: Chongqing University
#542	Title: Rule-Based Collaborative Learning with Heterogeneous Local Learning Models Authors: Ying Pang, Haibo Zhang, Jeremiah D. Deng, Lizhi Peng, Fei Teng Presenter: Ying Pang Affiliation: University of Otago
#257	Title: Data Removal from an AUC Optimization Model Authors: Jie Li, Jun-Qi Guo, Wei Gao Presenter: Jie Li Affiliation: Nanjing University
#355	Title: Adaptive Feature Generation for Online Continual Learning from Imbalanced Data Authors: Yingchun Jian, Jinfeng Yi, Lijun Zhang Presenter: Yingchun Jian Affiliation: Nanjing University



#178	Title: A Novel Semi-supervised Neural Network for recognizing Parkinson's Disease Authors: Zhehao Zhang, Xiaobo Zhang, Dengmin Wen, Lilan Peng, Yuxin Zhou Presenter: Zhehao Zhang Affiliation: Southwest Jiaotong University
#505	Title: Dual-State Knowledge Tracing Model with Mutual Information Maximization Authors: Haodong Meng, Changzhi Chen, Hongyu Yi, Xiaofeng He Presenter: Haodong Meng Affiliation: East China Normal University

Session | T3 – Foundation

May 17/Tuesday

Room: Diamond Ballroom / 2F | 璀璨宴会厅

Zoom ID: 815 9034 3067

Time: 16:10-18:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Fan Min

#361	Title: Robust and Provable Guarantees for Sparse Random Embeddings Authors: Maciej Skorski Alessandro Temperoni Martin Theobald Presenter: Alessandro Temperoni Affiliation: University of Luxembourg
#561	Title: Safe Offline Reinforcement Learning Through Hierarchical Policies Authors: Shaofan Liu, Shiliang Sun Presenter: Shaofan Liu Affiliation: East China Normal University
#413	Title: Hard Negative Sample Mining for Contrastive Representation in Reinforcement Learning Authors: Qihang Chen, Dayang Liang, Yunlong Liu Presenter: Qihang Chen Affiliation: Xiamen University
#386	Title: Transferable Interpolated Adversarial Attack with Random-Layer Mixup Authors: Size Ma, Keji Han, Xianzhong Long, Yun Li Presenter: Size Ma Affiliation: Nanjing University of Posts and Telecommunications
#83	Title: Self-paced Safe Co-training for Regression Authors: Fan Min, Yu Li, Liyan Liu Presenter: Fan Min Affiliation: Southwest Petroleum University



#134	Title: Attention-to-Embedding Framework for Multi-Instance Learning Authors: Mei Yang, Yu-Xuan Zhang, Mao Ye, Fan Min Presenter: Mei Yang Affiliation: Southwest Petroleum University
#279	Title: Hypersphere Neighborhood Rough Set for Rapid Attribute Reduction Authors: Yu Fang, Xue-Mei Cao, Xin Wang, Fan Min Presenter: Yu Fang Affiliation: Southwest Petroleum University
#141	Title: Multi-Instance Embedding Learning through High-Level Instance Selection Authors: Mei Yang, Wen-Xi Zeng, Fan Min Presenter: Wen-Xi Zeng Affiliation: Southwest Petroleum University

Session | V1- Data Science

May 17/Tuesday

Zoom ID: 829 6485 3183

Time: 10:30-12:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Hongjun Wang

#486	Title: A New Skeleton-Neural DAG Learning Approach Authors: Yiwen Cao, Kui Yu, Xiaoling Huang, Yujie Wang Presenter: Yiwen Cao Affiliation: Hefei University of Technology
#104	Title: Predicting Abnormal Events in Urban Rail Transit Systems with Multivariate Point Process Authors: Xiaoyun Mo, Mingqian Li, Mo Li Presenter: Xiaoyun Mo Affiliation: Nanyang Technological University
#653	Title: Simulate Human Thinking: Cognitive Knowledge Graph Reasoning for Complex Question Answering Authors: Hong Zhao, Yao Fu, Weihao Jiang, Shiliang Pu, Xiaoyu Cai Presenter: Hong Zhao Affiliation: Hikvision Research Institute, Hikvision
#654	Title: Separate then Constrain: A Hierarchical Network for End-to-End Triples Extraction Authors: Huizhao Wang, Yao Fu, Linghui Hu, Weihao Jiang, Shiliang Pu Presenter: Huizhao Wang Affiliation: Hikvision Research Institute, Hikvision



#323	Title: Semantics-Guided Disentangled Learning for Recommendation Authors: Dianer Yu, Qian Li, Xiangmeng Wang, Zhichao Wang, Yanan Cao, Guandong Xu Presenter: Dianer Yu Affiliation: University of Technology Sydney
#221	Title: Node Information Awareness Pooling for Graph Representation Learning Authors: Chuan Sun, Feihu Huang, Jian Peng Presenter: Chuan Sun Affiliation: Sichuan University
#705	Title: A Hybrid Semantic-Topic Co-encoding Network for Social Emotion Classification Authors: Lu Dai, Bang Wang, Wei Xiang, Minghua Xu, Han Xu Presenter: Lu Dai Affiliation: Huazhong University of Science and Technology

Session | V2 – Foundation

May 17/Tuesday

Zoom ID: 822 3121 1337

Time: 10:30-12:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Zhipeng Luo

#80	Title: Quantum Entanglement Inspired Correlation Learning for Classification Authors: Junwei Zhang, Zhao Li, Juan Wang, Yinghui Wang, Shichang Hu, Jie Xiao, Zhaolin Li Presenter: Junwei Zhang Affiliation: Tianjin University
#41	Title: Convergence and Applications of Alternating Direction Method of Multipliers on the Multi-convex Problems Authors: Junxiang Wang, Liang Zhao Presenter: Junxiang Wang Affiliation: Emory University
#568	Title: Leveraged Mel spectrograms using Harmonic and Percussive Components in Speech Emotion Recognition Authors: David Hason Rudd, Huan Huo, Guandong Xu Presenter: David Hason Rudd Affiliation: University of Technology Sydney
#162	Title: High Average-Utility Itemset Sampling under Length Constraints Authors: Lamine Diop Presenter: Lamine Diop Affiliation: University of Tours



#402	Title: Assessing Classifier Fairness with Collider Bias Authors: Zhenlong Xu, Ziqi Xu, Jixue Liu, Debo Cheng, Jiuyong Li, Lin Liu, Ke Wang Presenter: Ziqi Xu Affiliation: University of South Australia
#227	Title: Divide and Imitate: Multi-Cluster Identification and Mitigation of Selection Bias Authors: Katharina Dost, Hamish Duncanson, Ioannis Ziogas, Patricia Riddle and Joerg Wicker Presenter: Katharina Dost Affiliation: University of Auckland
#498	Title: Fact Aware Multi-Task Learning for Text Coherence Modeling Authors: Tushar Abhishek, Daksh Rawat, Manish Gupta, Vasudeva Varma Presenter: Tushar Abhishek Affiliation: IIIT Hyderabad

Session | V3 – Application

May 17/Tuesday

Zoom ID: 829 6485 3183

Time: 13:30-15:50 (Duration for Each Presentation: 17 minutes)

Session Chair: Chongshou Li

#405	Title: S^2QL: Retrieval Augmented Zero-shot Question Answering over Knowledge Graph Authors: Daoguang Zan, Sirui Wang, Hongzhi Zhang, Yuanmeng Yan, Wei Wu, Bei Guan, Yongji Wang Presenter: Daoguang Zan Affiliation: Institute of Software, Chinese Academy of Sciences; University of Chinese Academy of Sciences
#133	Title: A Novel Bayesian Deep Learning Approach to the Downscaling of Wind Speed with Uncertainty Quantification Authors: Firas Gerges, Michel C. Boufadel, Elie Bou-Zeid, Hani Nassif, Jason T. L. Wang Presenter: Firas Gerges Affiliation: New Jersey Institute of Technology
#196	Title: Recommending Personalized Interventions to Increase Employability of Disabled Jobseekers Authors: Ha Xuan Tran, Thuc Duy Le, Jiuyong Li, Lin Liu, Jixue Liu, Yanchang Zhao, Tony Waters Presenter: Ha Xuan Tran Affiliation: University of South Australia



#86	Title: Multicommunity Graph Convolution Networks with Decision Fusion for Personalized recommendation Authors: Shenghao Liu , Bang Wang, Bin Liu, Laurence T. Yang Presenter: Shenghao Liu Affiliation: Huazhong University of Science and Technology
#291	Title: A Two-stage Self-adaptive Model for Passenger Flow Prediction on Schedule-based Railway System Authors: Boyu Li, Ting Guo, Ruimin Li, Yang Wang, Amir Gandomi, Fang Chen Presenter: Boyu Li Affiliation: University of Technology Sydney
#159	Title: IDSGAN: Generative Adversarial Networks for Attack Generation against Intrusion Detection Authors: Zilong Lin, Yong Shi, Zhi Xue Presenter: Zilong Lin Affiliation: Indiana University Bloomington
#90	Title: Deep Learning for Prawn Farming: Forecasting and Anomaly Detection Authors: Joel Janek Dabrowski, Ashfaqur Rahman, Andrew Hellicar, Mashud Rana, Stuart Arnold Presenter: Joel Dabrowski Affiliation: CSIRO
#65	Title: NewsKVQA: Knowledge-Aware News Video Question Answering Authors: Pranay Gupta, Manish Gupta Presenter: Manish Gupta Affiliation: Microsoft

Session | V4 – Foundation

May 17/Tuesday

Zoom ID: 822 3121 1337

Time: 13:30-15:50 (Duration for Each Presentation: 17 minutes)

Session Chair: Laurence Park

#108	Title: Uniform Evaluation of Properties in Activity Recognition Authors: Seyed Mohammad Reza Modaresi, Aomar Osmani, Mohammadreza Razzazi, Abdelghani Chibani Presenter: Seyed Mohammad Reza Modaresi Affiliation: LIPN-CNRS, University Sorbonne Paris Nord
#36	Title: ENDASh: Embedding Neighbourhood Dissimilarity with Attribute Shuffling for Graph Anomaly Detection Authors: Qizhou Wang, Mahsa Salehi, Jia Shun Low, Wray Buntine, Christopher Leckie Presenter: Qizhou Wang Affiliation: Monash University

PAKDD2©22

#427	Title: Modeling IsA Relations via Box Structure for Knowledge Graph Embedding Authors: Yao Dong, Lei Wang, Ji Xiang, Kai Liu Presenter: Yao Dong Affiliation: Institute of Information Engineering, Chinese Academy of Sciences
#510	Title: Open Set Recognition for Time Series Classification Authors: Tolga Akar, Thorben Werner, Vijaya Krishna Yalavarthi, Lars Schmidt-Thieme Presenter: Tolga Akar Affiliation: Technical University of Berlin
#33	Title: Text2Chart: A Multi-Staged Chart Generatorfrom Natural Language Text Authors: Md. Mahinur Rashid, Hasin Kawsar Jahan, Annysha Huzzat, Riyasaat Ahmed Rahul, Tamim Bin Zakir, Farhana Meem, Md. Saddam Hossain Mukta, and Swakkhar Shatabda Presenter: Swakkhar Shatabda Affiliation: United International University
#627	Title: Cross-Lingual Product Retrieval in E-Commerce Search Authors: Wenya Zhu, Xiaoyu Lv, Baosong Yang, Yinghua Zhang, Xu Yong, Linlong Xu, Yinfu Feng, Haibo Zhang, Qing Da, Anxiang Zeng Presenter: Yinfu Feng Affiliation: Alibaba Group
#346	Title: Modelling Zeros in Blockmodelling Show abstract Authors: Laurence A. F. Park, Mohadeseh Ganji, Emir Demirovic, Jeffrey Chan, Peter Stuckey, James Bailey, Christopher Leckie, Rao Kotagiri Presenter: Laurence Park Affiliation: Western Sydney University
#388	Title: Deep Depression Prediction on Longitudinal Data via Joint Anomaly Ranking and Classification Authors: Guansong Pang, Ngoc Thien Anh Pham, Emma Baker, Rebecca Bentley, Anton Van Den Hengel Presenter: Guansong Pang Affiliation: Singapore Management University



Session | V5 – Data Science

May 17/Tuesday

Zoom ID: 829 6485 3183

Time: 16:10-18:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Chongshou Li

#85	Title: Emerging Scientific Topic Discovery by Finding Infrequent Synonymous Biterms Authors: Junfeng Wu, Guangyan Huang, Roozbeh Zarei, Jianxin Li, Guang-Li Huang, Hui Zheng, Jing He Presenter: Junfeng Wu Affiliation: Deakin University
#111	Title: Are Edge Weights in Summary Graphs Useful? - A Comparative Study Authors: Shinhwan Kang, Kyuhan Lee, Kijung Shin Presenter: Shinhwan Kang Affiliation: KAIST
#382	Title: Real-Time Skill Discovery in Intelligent Virtual Assistants Authors: Preeti Gopal Presenter: Preeti Gopal, Sunil Gupta, Santu Rana, Vuong Le, Trong Nguyen, and Svetha Venkatesh Affiliation: Deakin University
#563	Title: Joint Feature and Labeling Function Adaptation for Unsupervised Domain Adaptation Authors: Fengli Cui, Yinghao Chen, Yuntao Du, Yikang Cao, Chongjun Wang Presenter: Fengli Cui Affiliation: Nanjing University
#547	Title: Memory-Efficient Minimax Distance Measures Authors: Fazeleh Hoseini, Morteza Haghir Chehreghani Presenter: Fazeleh Hoseini Affiliation: Chalmers University of Technology
#170	Title: Coded Hate Speech Detection via Contextual Information Authors: Depeng Xu, Shuhan Yuan, Yueyang Wang, Angela Uchechukwu Nwude, Lu Zhang, Anna Zajicek and Xintao Wu Presenter: Xintao Wu Affiliation: University of Arkansas
#49	Title: PGADA: Perturbation-Guided Adversarial Alignmentfor Few-shot Learning Under the Support-Query Shift Authors: Siyang Jiang, Wei Ding, Hsi-Wen Chen, Ming-Syan Chen Presenter: Hsi-Wen Chen Affiliation: NTU



Title: dK-Personalization: Publishing Network Statistics with Personalized

Differential Privacy

#236 Authors: Masooma Iftikhar, Qing Wang, Yang Li

Presenter: Masooma Iftikhar

Affiliation: The Australian National University

Session | V6 - Application

May 17/Tuesday

Zoom ID: 822 3121 1337

Time: 16:10-18:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Dun Liu

#528	Title: DHA: Product Title Generation with Discriminative Hierarchical Attention for E-commerce Authors: Wenya Zhu, Yinghua Zhang, Yu Zhang, Yuhang Zhou, Yinfu Feng, Qing Da, Anxiang Zeng, Yuxiang Wu Presenter: Yinfu Feng Affiliation: Alibaba Group
#358	Title: Detail Perception Network for Semantic Segmentation in Water Scenes Authors: Cuixiao Liang, Wenjie Cai, Shaowu Peng, Qiong Liu Presenter: Cuixiao Liang Affiliation: South China University of Technology
#311	Title: ToothCR: A Two-stage Completion and Reconstruction Approach on 3D Dental Model Authors: Haoyu Zhu, Xiuyi Jia, Changdong Zhang, Tingting Liu Presenter: Haoyu Zhu Affiliation: Nanjing University of Science and Technology
#534	Title: Multi-Channel Orthogonal Decomposition Attention Network for Sequential Recommendation Authors: Jia Guo, Wendi Ji, Jiahao Yuan, and Xiaoling Wang Presenter: Jia Guo Affiliation: East China Normal University
#549	Title: Parametric Bandits for Search Engine Marketing Optimisation Authors: Marco Gigli, Fabio Stella Presenter: Marco Gigli Affiliation: University of Milano-Bicocca
#312	Title: BaDumTss: Multi-task Learning for Beatbox Transcription Authors: Priya Mehta, Meet Maheshwari, Brihi Joshi, Tanmoy Chakraborty Presenter: Meet Maheshwari Affiliation: IIIT Delhi



#502	Title: Extreme Multi-Label Classification with Hierarchical Multi-task for Product Attribute Identification Authors: Jun Zhang, Menqian Cai, Chenyu Zhao, Xiaowei Zhang, Zhiqian Zhang, Haiheng Chen, Sulong Xu Presenter: Jun Zhang Affiliation: JD.com
#280	Title: An Incentive Dispatch Algorithm for Utilization-Perfect EV Charging Management Authors: Lo Pang-Yun Ting, Po-Hui Wu, Hsiu-Ying Chung, and Kun-Ta Chuang Presenter: Lo Pang-Yun Ting Affiliation: NCKU

Session | T4 – Data Science

May 18/Wednesday

Room: Diamond Ballroom / 2F | 璀璨宴会厅

Zoom ID: 815 9034 3067

Time: 09:30-10:40 (Duration for Each Presentation: 17 minutes)

Session Chair: Zhi Xu

#316	Title: Distributed Differentially Private Ranking Aggregation Authors: Baobao Song, Qiujun Lan, Yang Li and Gang Li Presenter: Baobao Song Affiliation: Hunan universally
#671	Title: Structure-aware Reasoning for Knowledge Base Question Answering Authors: Lu Ma, Peng Zhang, Xi Zhu, Dan Luo, Bin Wang Presenter: Lu Ma Affiliation: Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy Sciences
#242	Title: Residual Vector Product Quantization for Approximate Nearest Neighbor search Authors: Zhi Xu, Lushuai Niu, Ruimin Meng, Longyang Zhao, and Jianqiu Ji Presenter: Lushuai Niu Affiliation: Guilin University of Electronic Technology
#185	Title: AutoTransformer: Automatic Transformer Architecture Design for Time Series Classification Authors: Yankun Ren, Longfei Li, Xinxing Yang, Jun Zhou Presenter: Yankun Ren Affiliation: Ant Group



Session | T5 – Foundation

May 18/Wednesday

Room: Diamond Ballroom / 2F | 璀璨宴会厅

Zoom ID: 815 9034 3067

Time: 11:00-12:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Chongshou Li

#602	Title: Neural Topic Modeling with Gaussian Mixture Model and Householder Flow Authors: Cangqi Zhou, Sunyue Xu, Hao Ban, Jing Zhang Presenter: Cangqi Zhou Affiliation: Nanjing University of Science and Technology
#288	Title: A Novel Clustering Algorithm with Dynamic Boundary Extraction Strategy Based on Local Gravitation Authors: Jiangmei Luo, Qingsheng Zhu, Junnan Li, Dongdong Cheng, Mingqiang Zhou Presenter: Jiangmei Luo Affiliation: Chongqing University
#465	Title: An GNN-Enhanced Game Bot Detection Model for MMORPGs Authors: Xianyang Qi, Jiashu Pu, Shiwei Zhao, Runze Wu, Jianrong Tao Presenter: Xianyang Qi Affiliation: Fuxi Al Lab, NetEase Games
#70	Title: Prototypical Classifier for Robust Class-Imbalanced Learning Authors: Tong Wei, Jiang-Xin Shi, Yu-Feng Li, Min-Ling Zhang Presenter: Tong Wei Affiliation: Nanjing University
#708	Title: Improving Energy-based Out-of-distribution Detection by Sparsity Regularization Authors: Qichao Chen; Wenjie Jiang; Kuan Li; Yi Wang Presenter: Qichao Chen Affiliation: Dongguan University of Technology



Session | V7 – Data Science

May 18/Wednesday

Zoom ID: 829 6485 3183

Time: 09:30-10:40 (Duration for Each Presentation: 17 minutes)

Session Chair: Mangesh Bendre

ONLINE

#198	Title: Knowledge Lock: Overcoming Catastrophic Forgetting in Federated Learning Authors: Guoyizhe Wei, Xiu Li Presenter: Guoyizhe Wei Affiliation: Tsinghua University
#349	Title: Multi-Task Knowledge Graph Representations via Residual Functions Authors: Adit Krishnan, Mahashweta Das, Mangesh Bendre, Fei Wang, Hao Yang, Hari Sundaram Presenter: Mangesh Bendre Affiliation: Visa Research
#529	Title: Multi-Granularity Evolution Network for Dynamic Link Prediction Authors: Yi Yang, Xiaoyan Gu, Haihui Fan, Bo Li, Weiping Wang Presenter: Yi Yang Affiliation: Institute of Information Engineering, Chinese Academy of Sciences
#207	Title: Aspect-based Sentiment Analysis through EDU-level Attentions Authors: Ting Lin and Aixin Sun, Yequan Wang Presenter: Ting Lin Affiliation: Nanyang Technological University

Session | V8 - Foundation

May 18/Wednesday

Zoom ID: 822 3121 1337

Time: 09:30-10:40 (Duration for Each Presentation: 17 minutes)

Session Chair: Zhipeng Luo

	Title: DeepPAMM: Deep Piecewise Exponential Additive Mixed Models for
	Complex Hazard Structures in Survival Analysis
#400	Authors: Philipp Kopper, Simon Wiegrebe, Bernd Bischl, Andreas Bender, David
#400	Rügamer
	Presenter: Philipp Kopper
	Affiliation: LMU Munich



#417	Title: Reduction of the Position Bias via Multi-Level Learning for Activity Recognition Authors: Aomar Osmani; Massinissa Hamidi Presenter: Massinissa Hamidi Affiliation: LIPN-CNRS, University Sorbonne Paris Nord
#638	Title: An adaptable indexing pipeline for enriching meta information of datasets from heterogeneous repositories Authors: Siamak Farshidi and Zhiming Zhao Presenter: Siamak Farshidi Affiliation: University of Amsterdam
#675	Title: SMITH: A Self-Supervised Downstream-aware Framework for Missing Testing Data Handling Authors: Chih-Chun Yang, Cheng-Te Li, and Shou-De Lin Presenter: Chih-Chun Yang Affiliation: NTU

Session | V9 - Data Science

May 18/Wednesday

Zoom ID: 829 6485 3183

Time: 11:00-12:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Hao Niu

#156	Title: Mu2ReST: Multi-Resolution Recursive Spatio-Temporal Transformer for Long-Term Prediction Authors: Hao Niu, Chuizheng Meng, Defu Cao, Guillaume Habault, Roberto Legaspi, Shinya Wada, Chihiro Ono, Yan Liu Presenter: Hao Niu Affiliation: KDDI Research, Inc.
#408	Title: Few-shot knowledge graph entity typing Authors: Guozhen Zhu, Zhongbao Zhang, Sen Su Presenter: Guozhen Zhu Affiliation: Beijing University of Posts and Telecommunications
#604	Title: Graph Multi-Head Convolution for Spatio-Temporal Attention in Origin Destination Tensor Prediction Authors: Manish Bhanu, Rahul Kumar, Saswata Roy, João Mendes-Moreira, Joydeep Chandra Presenter: Manish Bhanu Affiliation: Indian Institute of Technology Patna



#356	Title: Order-Aware Graph Neural Network for Sequential Recommendation Authors: Xinlei Zhang, Wendi Ji, Jiahao Yuan, Xiaoling Wang Presenter: Xinlei Zhang Affiliation: East China Normal University
	Title: ADAM: An Attentional Data Augmentation Method for Extreme Multi-Label Text Classification Authors: Jiaxin Zhang, Jie Liu, Shaowei Chen, Shaoxin Lin, Bingquan Wang,
#181	Shanpeng Wang Presenter: Jiaxin Zhang
	Affiliation: Nankai University

Session | V10 – Foundation

May 18/Wednesday

Zoom ID: 822 3121 1337

Time: 11:00-12:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Guilin Zhao

#686	Title: Layer Adaptive Deep Neural Networks for Out-of-distribution Detection Authors: Haoliang Wang, Chen Zhao, Xujiang Zhao, Feng Chen Presenter: Haoliang Wang Affiliation: University of Texas at Dallas
#560	Title: Discretization inspired defence Algorithm against Adversarial Attacks of Tabular Data Authors: Jiahui Zhou, Nayyar Zaidi, Yishuo Zhang, Gang Li Presenter: Jiahui Zhou Affiliation: Xi'an Shiyou University
#348	Title: Towards Better Generalization for Neural Network- based SAT Solvers Authors: Chenhao Zhang, Yanjun Zhang, Jeff Mao, Weitong Chen, Lin Yue Guangdong Bai, Miao Xu Presenter: Chenhao Zhang Affiliation: The University of Queensland
#591	Title: SelectAug: A Data Augmentation Method for Distracted Driving Detection Authors: Yuan Li, Wei Mi, Jingguo Ge, Jingyuan Hu, Hui Li, Daoqing Zhang,Tong Li Presenter: Yuan Li Affiliation: Institute of Information Engineering, Chinese Academy of Sciences China School of Cyber Security, University of Chinese Academy of Sciences
#624	Title: Contrastive Attributed Network Anomaly Detection with Data Augmentation Authors: Zhiming Xu, Xiao Huang, Yue Zhao, Yushun Dong, Jundong Li Presenter: Zhiming Xu Affiliation: University of Virginia



Session | V11 – Data Science

May 19/Thursday

Zoom ID: 815 9034 3067

Time: 09:00-10:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Yan Yang

#428	Title: Improving Entity Disambiguation using Knowledge Graph Regularization Authors: Zhi Rui Tam, Yi-Lun Wu, Hong-Han Shuai Presenter: Hong-Han Shuai Affiliation: NYCU
#651	Title: Instance-guided Multi-modal Fake News Detection with Dynamic Intra- and Inter-Modality Fusion Authors: Jie Wang, Yan Yang, Keyu Liu, Peng Xie, Xiaorong Liu Presenter: Jie Wang Affiliation: Southwest Jiaotong University
#67	Title: Auxiliary Local Variables for Improving Regularization/prior Approach in Continual Learning Authors: Linh Ngo Van, Nam Le Hai, Hoang Pham, Khoat Than Presenter: Linh Ngo Van Affiliation: Hanoi University of Science and Technology
#616	Title: Smooth Perturbations for Time Series Adversarial Attacks Authors: Gautier Pialla, Hassan Ismail Fawaz, Maxime Devanne, Jonathan Weber, Lhassane Idoumghar, Pierre-Alain Muller, Christoph Bergmeir, Daniel Schmidt, Geoffrey Webb, Germain Forestier Presenter: Gautier Pialla Affiliation: Université de Haute-Alsace
#175	Title: Reducing catastrophic forgetting in neural networks via Gaussian mixture approximation Authors: Hoang Phan Viet, Anh Phan Tuan, Son Nguyen, Linh Ngo Van, Khoat Than Presenter: Linh Ngo Van Affiliation: Hanoi University of Science and Technology



Session | V12 – Data Science

May 19/Thursday

Zoom ID: 829 6485 3183

Time: 09:00-10:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Fengmao Lv

#541	Title: A Two-Tower Spatial-temporal Graph Neural Network for Traffic Speed Prediction Authors: Yansong Shen, Lin Li, Qing Xie, Xin Li, GuanDong Xu Presenter: Yansong Shen Affiliation: Wuhan University of Technology
#481	Title: Partially Relaxed Masks for Knowledge Transfer without Forgetting in Continual Learning Authors: Tatsuya Konishi, Mori Kurokawa, Chihiro Ono, Zixuan Ke, Gyuhak Kim, Bing Liu Presenter: Tatsuya Konishi Affiliation: KDDI Research, Inc.
#610	Title: Evolution-based Online Automated Machine Learning Authors: Cedric Presenter: Cedric Kulbach Affiliation: FZI Research Center for Information Technology
#459	Title: PASTA: PArallel Spatio-Temporal Attention with Spatial Auto-correlation Gating for Fine-grained Crowd Flow Prediction Authors: Chung Park, Junui Hong, Cheonbok Park, Taesan Kim, Minsung Choi, Jaegul Choo Presenter: Chung Park Affiliation: SK Telecom and KAIST (Korea Advanced Institute of Science and Technology)
#669	Title: Sparse Imbalanced Drug-Target Interaction Prediction via Heterogeneous Data Augmentation and Node Similarity Authors: Runze Wang, Zehua Zhang, Yueqin Zhang, Zhongyuan Jiang, Shilin Sun, Chenwei Zhang Presenter: Runze Wang Affiliation: Taiyuan University of Technology



Session | V13 – Foundation

May 19/Thursday

Zoom ID: 822 3121 1337

Time: 09:00-10:30 (Duration for Each Presentation: 17 minutes)

Session Chair: Donghun Lee

#684	Title: Tlife-GDN: Detecting and Forecasting Spatio-Temporal Anomalies via Persistent Homology and Geometric Deep Learning Authors: Zhiwei Zhen, Yuzhou Chen, Ignacio Segovia-Dominguez and Yulia R. Gel Presenter: Zhiwei Zhen Affiliation: University of Texas at Dallas
#623	Title: Dynamic Topic-Noise Models for Social Media Authors: Rob Churchill, Lisa Singh Presenter: Rob Churchill Affiliation: Georgetown University
#132	Title: Effect of Different Encodings and Distance Functions on Quantum Instance-based Classifiers Authors: Alessandro Berti, Anna Bernasconi, Gianna M. Del Corso, Riccardo Guidotti Presenter: Alessandro Berti Affiliation: University of Pisa
#657	Title: FLiB: Fair Link Prediction in Bipartite Network Authors: Piyush Kansal, Nitish Kumar, Sangam Verma, Karamjit Singh, Pranav Poduval Presenter: Nitish Kumar Affiliation: MASTERCARD
#495	Title: Online Learning with Regularized Knowledge Gradients Authors: Donghun Lee, Warren B. Powell Presenter: Donghun Lee Affiliation: Korea University



Session | V14 - Data Science

May 19/Thursday

Zoom ID: 815 9034 3067

Time: 10:50-12:00 (Duration for Each Presentation: 17 minutes)

Session Chair: Zhiguo Long

ONLINE

#680	Title: Detecting Anchors' Opinion in Hinghlish News Delivery Authors: Siddharth Sadhwani, Nishant Grover, Md Shad Akhtar, Tanmoy Chakraborty Presenter: Siddharth Sadhwani Affiliation: Indraprastha Institute of Information Technology - Delhi
#632	Title: Misleading Inference Generation via Proximal Policy Optimization Authors: Hsien-Yung Peng, Ho-Lam Chung, Ying-Hong Chan, Yao-Chung Fan Presenter: Yao-Chung Fan Affiliation: NCHU
#567	Title: Protoformer: Embedding Prototypes for Transformers Authors: Ashkan Farhangi Ning Sui Nan Hua Haiyan Bai Arthur Huang Zhishan Guo Presenter: Ning Sui Affiliation: University of Central Florida
#214	Title: Interconnected Neural Linear Contextual Bandits with UCB Exploration Authors: Yang Chen, Miao Xie, Jiamou Liu, Kaiqi Zhao Presenter: Yang Chen Affiliation: The University of Auckland

Session | V15 – Application

May 19/Thursday

Zoom ID: 829 6485 3183

Time: 10:50-12:00 (Duration for Each Presentation: 17 minutes)

Session Chair: Dawei Zhan

#587

ONLINE

Title: User Incentive Based Bike-sharing Dispatching Strategy

Authors: Bing Shi, Zhaoxiang Song, Xizi Huang, and Jianqiao Xu

Presenter: Zhaoxiang Song

Affiliation: Wuhan University of Technology



#205	Title: Estimating Skill Proficiency from Resumes Authors: Anindita Sinha Banerjee, Sachin Pawar, Girish K. Palshikar, Devavrat Thosar, Jyoti Bhat, Payodhi Mandloi Presenter: Sachin Pawar Affiliation: TCS Research, Tata Consultancy Services
#539	Title: Insomnia Disorder Detection using EEG Sleep Trajectories Authors: Stephen McCloskey, Bryn Jeffries, Irena Koprinska, Christopher Gordon, and Ronald R. Grunstein Presenter: Stephen McCloskey Affiliation: The University of Sydney

Session | V16 – Application

May 19/Thursday

Zoom ID: 822 3121 1337

Time: 10:50-12:00 (Duration for Each Presentation: 17 minutes)

Session Chair: Fengmao Lv

#512	Title: Exploiting Spatial Attention and Contextual Information for Document Image Segmentation Authors: Yuman Sang, Yifeng Zeng, Ruiying Liu, Fan Yang, Zhangrui Yao, Yinghui Pan Presenter: Yuman Sang Affiliation: Xiamen University
#370	Title: Mental Health Treatments Using an Explainable Adaptive Clustering Model Authors: Usman Ahmed, Jerry Chun-Wei Lin, Gautam Srivastava Presenter: Usman Ahmed Affiliation: Western Norway University of Applied Sciences
#601	Title: ALBIF: Active Learning With BandIt Feedbacks Authors: Mudit Agarwal, Naresh Manwani Presenter: Mudit Agarwal Affiliation: IIIT Hyderabad



MEMO