KDD2019

ANCHORAGE, ALASKA

AUGUST 4–8, 2019

Dena'ina Convention Center and
William Egan Convention Center

25TH ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING

The annual KDD conference is the premier interdisciplinary conference bringing together researchers and practitioners from data science, data mining, knowledge discovery, large-scale data analytics, and big data.

PROGRAM

WWW.KDD.ORG/KDD2019/

Association for Computing Machinery

TABLE OF **CONTENTS**

4 Agenda at a Glance

10 General Chairs' Welcome Message

11 Program Chairs' Welcome Message

12 Program Highlights
- Lecture–Style Tutorial Program
- Workshop Program
- Earth Day Program
- Deep Learning Day Program
- Health Day Program
- KDD Cup Day Program
- Project Showcase Program
- Keynotes
- Keynote Panel
- Women's Lunch
- Social Impact Workshop
- Hands-On Tutorial Program
- Applied Data Science Invited Talks
- ADS and Research Track Oral Presentations
- Awards

36 KDD 2019 Organizing Committee

37 Sponsors & Supporters

KDD2019

www.kdd.org/kdd2019/

AGENDA
AT A GLANCE

KDD 2019: Sunday, August 4 (TUTORIAL DAY)

7:00AM - 5:00PM	**KDD 2019 Registration** – Tikahtnu Foyer- Level 3, Dena'ina
8:00AM - 5:00PM	**T1: Learning From Networks: Algorithms, Theory, & Applications (FULL DAY)** – Summit 1- Ground Level, Egan
8:00AM - 12:00PM	**T2: Data Integration and Machine Learning: A Natural Synergy** – Kahtnu 2- Level 2, Dena'ina
8:00AM - 12:00PM	**T3: Mining and model understanding on medical data** – Tubughnenq 3- Level 2, Dena'ina
8:00AM - 12:00PM	**T4: Deep Reinforcement Learning with Applications in Transportation** – Tubughnenq 4- Level 2, Dena'ina
8:00AM - 12:00PM	**T5: Gold Panning from the Mess: Rare Category Exploration, Exposition, Representation and Interpretation** – Kahtnu 1- Level 2, Dena'ina
8:00AM - 12:00PM	**T6: Social User Interest Mining: Methods and Applications** – Tubughnenq 5- Level 2, Dena'ina
8:00AM - 12:00PM	**T7: Fairness-Aware Machine Learning: Practical Challenges and Lessons Learned** – Summit 11- Ground Level, Egan
8:00AM - 12:00PM	**T8: Modern MDL meets Data Mining -- Insights, Theory, and Practice** – Summit 2- Ground Level, Egan
8:00AM - 12:00PM	**T9: Adaptive Influence Maximization** – Summit 3- Ground Level, Egan
8:00AM - 12:00PM	**T10: Hypothesis Testing and Statistically-sound Pattern Mining** – Summit 4- Ground Level, Egan
8:00AM - 12:00PM	**T11: Fake News Research: Theories, Detection Strategies, and Open Problems** – Summit 5- Ground Level, Egan
8:00AM - 12:00PM	**T12: Are You My Neighbor? Bringing Order to Neighbor Computing Problems** – Summit 6- Ground Level, Egan
8:00AM - 12:00PM	**T13: Mining temporal networks** – Summit 7- Ground Level, Egan
8:00AM - 12:00PM	**T14: Modeling and Applications for Temporal Point Processes** – Summit 8- Ground Level, Egan
8:00AM - 12:00PM	**T15: Challenges, Best Practices and Pitfalls in Evaluating Results of Online Controlled Experiments** – Summit 9- Ground Level, Egan
9:30AM - 10:00AM	**KDD Coffee Break** – Lobby Foyers- Dena'ina and Egan
12:00PM-1:00PM	**Boxed Lunches** – Lobby Foyers- Dena'ina and Egan
1:00PM - 5:00PM	**T1: Learning From Networks: Algorithms, Theory, & Applications (FULL DAY)** – Summit 1- Ground Level, Egan
1:00PM - 5:00PM	**T16: Foundations of large-scale sequential experimentation** – Kahtnu 1- Level 2, Dena'ina
1:00PM - 5:00PM	**T17: Constructing and Mining Heterogeneous Information Networks from Massive Text** – Kahtnu 2- Level 2, Dena'ina
1:00PM - 5:00PM	**T18: Deep Bayesian Mining, Learning and Understanding** – Tubughnenq 3- Level 2, Dena'ina
1:00PM - 5:00PM	**T19: Incompleteness in Networks: Biases, Skewed Results, and Some Solutions** – Tubughnenq 4- Level 2, Dena'ina
1:00PM - 5:00PM	**T20: Optimize the Wisdom of the Crowd: Inference, Learning, and Teaching** – Tubughnenq 5- Level 2, Dena'ina
1:00PM - 5:00PM	**T21. Interpretable Knowledge Discovery Reinforced by Visual Methods** – Summit 11- Ground Level, Egan
1:00PM - 5:00PM	**T22. Explainable AI in Industry** – Summit 2- Ground Level, Egan
1:00PM - 5:00PM	**T23. Advances in Cost-sensitive Multiclass and Multilabel Classification** – Summit 3- Ground Level, Egan
1:00PM - 5:00PM	**T24. Recent Progress in Zeroth Order Optimization and Its Applications to Adversarial Robustness in Data Mining and Machine Learning** – Summit 4- Ground Level, Egan
1:00PM - 5:00PM	**T25. Forecasting Big Time Series: Theory and Practice** – Summit 5- Ground Level, Egan
1:00PM - 5:00PM	**T26. Deep Natural Language Processing for Search and Recommender Systems** – Summit 6- Ground Level, Egan
1:00PM - 5:00PM	**T27. Spatio-temporal event forecasting and precursor identification** – Summit 7- Ground Level, Egan
1:00PM - 5:00PM	**T28. Statistical Mechanics Methods for Discovering Knowledge from Production-Scale Neural Networks** – Summit 8- Ground Level, Egan
1:00PM - 5:00PM	**T29. Data Mining Methods for Drug Discovery and Development**– Summit 9- Ground Level
2:30PM - 3:00PM	**KDD Coffee Break** – Lobby Foyers- Dena'ina and Egan

KDD 2019: Monday, August 5 (WORKSHOP DAY)

7:00AM - 5:00PM	**KDD 2019 Registration** – Tikahtnu Foyer- Level 3, Dena'ina
8:00AM – 5:00PM	**Sponsor Interview Room** – Kodiak Boardroom- Marriott Hotel
8:00AM - 5:00PM	**W1: The 5th Workshop in Mining and Learning from Time Series** – Tubughnenq 3- Level 2, Dena'ina
8:00AM – 5:00PM	**W2: AdKDD 2019 Workshop** – Kahtnu 2- Level 2, Dena'ina
8:00AM - 5:00PM	**W3: 15th International Workshop on Mining and Learning with Graphs (MLG)** – Kahtnu 1- Level 2, Dena'ina
8:00AM – 12:00PM	**Deep Learning Day (see page 16 for full schedule)** – La Perouse- Street Level, Egan
8:00AM – 12:00PM	**Health Day (see page 17 for full schedule)** – Cook- Street Level, Egan
8:00AM - 12:00PM	**W4: [Earth Day]: Fragile Earth: Theory Guided Data Science to Enhance Scientific Discovery** – Summit 3- Ground Level, Egan
8:00AM - 12:00PM	**W5: [Earth Day]: Data Mining and AI for Conservation**– Summit 6- Ground Level, Egan
8:00AM - 12:00PM	**W6: [Earth Day]: The 8th KDD International Workshop on Urban Computing** – Summit 1- Ground Level, Egan
8:00AM - 12:00PM	**W7: The 8th International Workshop on Parallel and Distributed Computing for Large-Scale Machine Learning and Big Data Analytics** – Tubughnenq 4- Level 2, Dena'ina
8:00AM - 12:00PM	**W8: The 1st International Workshop on Artificial Intelligence of Things** – Summit 14- Ground Level, Egan
8:00AM - 12:00PM	**W9: 2nd Workshop on Offline and Online Evaluation of Interactive Systems** – Summit 13- Ground Level, Egan
8:00AM - 12:00PM	**W10: KDD 2019 Workshop on Causal Discovery (CD2019)** – Summit 11- Ground Level, Egan
8:00AM - 12:00PM	**W11: The First International Workshop on Smart Data for Blockchain and Distributed Ledger (SDBD'19)** – Summit 12- Ground Level, Egan
8:00AM - 12:00PM	**W12: Tensor Methods for Emerging Data Science Challenges** – Summit 9- Ground Level, Egan

8:00AM - 12:00PM	**W13: 2nd KDD Workshop on Anomaly Detection in Finance** – Summit 8- Ground Level, Egan
8:00AM - 12:00PM	**W14: Explainable AI/ML (XAI) for Accountability, Fairness, and Transparency** – Summit 7- Ground Level, Egan
8:00AM - 12:00PM	**W15: AI for Fashion: The Fourth International Workshop on Fashion and KDD** – Summit 5- Ground Level, Egan
8:00AM - 12:00PM	**W16: Data Collection, Curation and Labeling for Mining and Learning** – Summit 4- Ground Level, Egan
8:00AM - 12:00PM	**W17: Adversarial Learning Methods for Machine Learning and Data Mining** – Tubughnenq 5- Level 2, Dena'ina
8:00AM - 12:00PM	**W18: 2019 International Workshop on Talent and Management Computing (TMC'2019)** – Summit 2- Ground Level, Egan
9:30AM - 10:00AM	**KDD Coffee Break** – Lobby Foyers- Dena'ina and Egan
12:00PM – 1:00PM	**Boxed Lunches** – Lobby Foyers- Dena'ina and Egan
1:00PM – 5:15 PM	**Earth Day (see page 16 for full schedule)** – Arteaga- Streel Level, Egan
1:00PM -5:00PM	**W19: [Deep Learning Day]: Deep Learning on Graphs: Methods and Applications (DLG'19)** – Summit 10- Ground Level, Egan
1:00PM - 5:00PM	**W20: [Deep Learning Day]: Deep Reinforcement Learning for Knowledge Discovery** – Summit 1- Ground Level, Egan
1:00PM - 5:00PM	**W21: [Deep Learning Day]: 1st International Workshop on Deep Learning Practice for High-Dimensional Sparse Data with KDD 2019** – Summit 2- Ground Level, Egan
1:00PM - 5:00PM	**W22: [Deep Learning Day]: Deep Learning for Education (DL4Ed)** – Summit 14- Ground Level, Egan
1:00PM - 5:00PM	**W23: [Health Day]: [DsHealth 2019] 2019 KDD Workshop on Applied Data Science in Healthcare: Bridging the Gap Between Data and Knowledge** – Summit 3- Ground Level, Egan
1:00PM - 5:00PM	**W24: [Health Day]: epiDAMIK: Epidemiology Meets Data Mining and Knowledge Discovery** – Summit 7- Ground Level, Egan
1:00PM - 5:00PM	**W25: [Health Day]: 18th International Workshop on Data Mining in Bioinformatics (BIOKDD'19)** – Summit 13- Ground Level, Egan Summit 6- Ground Level, Egan
1:00PM - 5:00PM	**W26: Developing Standards on Definitions of Analytics Roles, Skillsets and Career Paths in the Data Science Industry** – Tubughnenq 5- Level 2, Dena'ina
1:00PM - 5:00PM	**W27: The First International Workshop on Intelligent Information Feed** – Summit 12- Ground Level, Egan
1:00PM - 5:00PM	**W28: The Third International Workshop on Automation in Machine Learning** – Summit 11- Ground Level, Egan
1:00PM - 5:00PM	**W29: BigMine-19: 8th International Workshop on Big Data, IoT Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications** – Tubughnenq 4- Level 2, Dena'ina
1:00PM - 5:00PM	**W30: Learning and Mining for Cyber Security (LEMINCS)** – Summit 13- Ground Level, Egan
1:00PM - 5:00PM	**W31: From Data Integration to Knowledge Graphs: Challenges and Experiences** – Summit 5- Ground Level, Egan
1:00PM - 5:00PM	**W32: Data-driven Intelligent Transportation (DIT'19)** – Summit 4- Ground Level, Egan
1:00PM - 5:00PM	**W33: Truth Discovery and Fact Checking: Theory and Practice** – Summit 8- Ground Level, Egan
1:00PM - 5:00PM	**W34: Eighth International Workshop on Issues of Sentiment Discovery and Opinion Mining(WISDOM'19)** – Summit 9- Ground Level, Egan
2:30PM - 3:00PM	**KDD Coffee Break** – Lobby Foyers- Dena'ina and Egan
5:30PM - 7:00PM	**KDD 2019 Opening Session** – Tikahtnu Ballroom- Level 3, Dena'ina
7:00PM - 9:30PM	**Poster Reception: Group 1** *Research Track Oral Papers, Applied Data Science Track Oral Papers, Deep Learning Day, and Health Day* – Idlughet Hall 3- Street Level, Dena'ina *A full list of poster titles and numbers can be found in the event app, and will be handed out at the door*

KDD 2019: Tuesday, August 6 (MAIN CONFERENCE DAY 1)

7:00AM - 5:00PM	**KDD 2019 Registration** – Tikahtnu Foyer- Level 3, Dena'ina
8:00AM - 9:30AM	**Opening Keynote: Peter Lee,** Corporate Vice President, Microsoft Healthcare**: The Unreasonable Effectiveness, and Difficulty, of Data in Healthcare** – Tikahtnu Ballroom- Level 3, Dena'ina
9:30AM - 6:00PM	**Sponsor Interview Room** – Kodiak Boardroom- Marriott Hotel
9:30AM - 6:00PM	**KDD Exhibit Hall** – Idlughet Hall- Street Level, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: Put Deep Learning to Work: A Practical Introduction to using Amazon Web Services** – Kahtnu- Level 2, Dena'ina
9:30AM – 12:30PM	**Hands-On Tutorial: Democratizing & Accelerating AI through Machine Learning** – Tubughnenq 3- Level 2, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: Concept to Code: Deep Neural Conversational System** – Tubughnenq 5- Level 2, Dena'ina
9:30AM - 10:00AM	**KDD Coffee Break** – Lobby Foyer- Dena'ina
10:00AM - 5:00PM	**KDD Cup Day (see page 18 for full schedule)** – Arteaga- Street Level, Egan
10:00AM - 12:00PM	**Applied Data Science Invited Talks (see page 23 for full schedule)** – Cook- Street Level, Egan
10:00AM - 12:00PM	**Research Track Oral Presentations: RT1 Neural Networks** – Summit 2- Ground Level, Egan
10:00AM – 12:00PM	**Research Track Oral Presentations: RT2 Analyzing Sequential and Temporal Data** – Summit 3- Ground Level, Egan
10:00AM – 12:00PM	**Research Track Oral Presentations: RT3 Algorithmic Techniques** – Summit 4- Ground Level, Egan
10:00AM – 12:00PM	**Applied Data Science Oral Presentations: ADS1 Auto-ML and Development Frameworks** – Summit 1- Ground Level, Egan
12:00PM – 1:30PM	**ADS and Research PC/SPC Luncheon (Invite Only)** – La Perouse- Street Level, Egan
12:00PM - 1:30PM	**KDD Lunch** – Idlughet Hall- Street Level, Dena'ina
1:30PM - 4:30PM	**Hands-On Tutorial: Cloud Based Data Science at the Speed of Thought Using RAPIDS: the open GPU Data Science Ecosystem** – Kahtnu- Level 2, Dena'ina
1:30PM - 4:30PM	**Hands-On Tutorial: Introduction to Computer Vision and Realtime Deep Learning-Based Object Detection** – Tubughnenq 3- Level 2, Dena'ina
1:30PM - 4:30PM	**Hands-On Tutorial: Declarative Text Understanding with SystemT** – Tubughnenq 5- Level 2, Dena'ina
1:30PM – 6:00PM	**Poster Blitz Sessions** – Tikahtnu Ballroom- Level 3, Dena'ina
1:30PM – 6:00PM	**Applied Data Science Invited Talks (see page 23 for full schedule)** – Cook- Street Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT4 Embeddings I** – Summit 2- Ground Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT5 Privacy and Policy Learning** – Summit 3- Ground Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT6 Network Science** – Summit 4- Ground Level, Egan
1:30PM - 3:30PM	**Applied Data Science Oral Presentations: ADS2 Language Models and Text Mining** – Summit 1- Ground Level, Egan
1:30PM - 3:30PM	**Applied Data Science Oral Presentations: ADS3 Urbanism and Mobility** – Summit 5/6- Ground Level, Egan
3:30PM - 4:00PM	**KDD Coffee Break** – Lobby Foyer- Dena'ina
4:00PM - 6:00PM	**Research Track Oral Presentations: RT7 Neural Networks** – Summit 2- Ground Level, Egan
4:00PM - 6:00PM	**Research Track Oral Presentations: RT8 Knowledge Extraction** – Summit 3- Ground Level, Egan
4:00PM - 6:00PM	**Research Track Oral Presentations: RT9 Mining In Emerging Applications I** – Summit 4- Ground Level, Egan
4:00PM - 6:00PM	**Applied Data Science Oral Presentations: ADS4 Real-time and Online** – Summit 1- Ground Level, Egan

4:00PM - 6:00PM	**Dissertation Award** – Summit 5- Ground Level, Egan
6:00PM - 7:00PM	**SIGKDD Business Meeting** – La Perouse- Street Level, Egan
7:00PM - 9:30PM	**Poster Reception: Group 2** ***Applied Data Science Track Poster Papers, Research Track Poster Papers, KDD Cup, Earth Day, Social Impact*** – Idlughet Hall 3- Street Level, Dena'ina *A full list of poster titles and numbers can be found in the event app, and will be handed out at the door*

KDD 2019: Wednesday, August 7 (MAIN CONFERENCE DAY 2)

8:00AM - 5:00PM	**KDD 2019 Registration** – Tikahtnu Foyer- Level 3, Dena'ina
8:00AM – 11:45AM	**KDD Social Impact Workshop (see page 21 for full schedule)** – **Tukatnu Ballroom- Level 3, Dena'ina**
8:00AM - 5:00PM	**Sponsor Interview Room** – Kodiak Boardroom- Marriott Hotel
8:00AM – 12:00PM	**KDD India Meeting** – Summit 5/6- Ground Level, Egan
8:30AM – 3:30PM	**KDD Project Showcase (see page 18 for full schedule)** – Idlughet Hall 3- Street Level, Dena'ina
9:30AM - 6:00PM	**KDD Exhibit Hall** – Idlughet Hall- Street Level, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: Deep Learning for NLP with TensorFlow** – Kahtnu- Level 2, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: Deep Learning at Scale on Databricks** – Tubughnenq 3- Level 2, Dena'ina
9:30AM – 12:30PM	**Hands-On Tutorial: Deep Learning for Time Series Forecasting** – Tubughnenq 5- Level 2, Dena'ina
9:30AM - 10:00AM	**KDD Coffee Break** – Lobby Foyer- Dena'ina
10:00AM - 12:00PM	**Applied Data Science Invited Talks (see page 23 for full schedule)** – Cook- Street Level, Egan
10:00AM - 12:00PM	**Research Track Oral Presentations: RT10 Embeddings I** – Summit 2- Ground Level, Egan
10:00AM - 12:00PM	**Research Track Oral Presentations: RT11 Clustering and Visualization** – Summit 3- Ground Level, Egan
10:00AM – 12:00PM	**Research Track Oral Presentations: RT12 Recommender Systems I** – Summit 4- Ground Level, Egan
10:00AM - 12:00PM	**Applied Data Science Oral Presentations: ADS5 Scalability and Novel Applications** – Summit 1- Ground Level, Egan
12:00PM - 1:30PM	**KDD Lunch** – Idlughet Hall- Street Level, Dena'ina
12:00PM - 1:30PM	**KDD Women's Lunch (Ticket Required)** – Arteaga- Street Level, Egan
1:30PM - 4:30PM	**Hands-On Tutorial: Introduction to Kubeflow Pipelines** – Kahtnu- Level 2, Dena'ina
1:30PM - 4:30PM	**Hands-On Tutorial: Deep Learning at Scale on Databricks** – Tubughnenq 3- Level 2, Dena'ina
1:30PM - 4:30PM	**Hands-On Tutorial: Learning Graph Neural Networks with Deep Graph Library** – Tubughnenq 5- Level 2, Dena'ina
1:30PM - 3:30PM	**Applied Data Science Invited Talks (see page 23 for full schedule)** – Cook- Street Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT13 Learning** – Summit 2- Ground Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT14 Anomaly Detection** – Summit 3- Ground Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT15 Mining in Emerging Applications II** – Summit 4- Ground Level, Egan
1:30PM - 3:30PM	**Applied Data Science Oral Presentations: ADS6 Environment and Sustainability** – Summit 1- Ground Level, Egan
3:30PM - 4:00PM	**KDD Coffee Break** – Lobby Foyer- Dena'ina
4:00PM - 6:00PM	**Plenary Keynote Panel: Why, What & How we Democratize Data Science (see page 20 for full details** – Tukatnu Ballroom- Level 3, Dena'ina
6:30PM – 10:00PM	**KDD 2019 25th Anniversary Celebration Dinner** – Tukatnu Ballroom, Foyer and Patio- Level 3, Dena'ina

KDD 2019: Thursday, August 8 (MAIN CONFERENCE DAY 3)

8:00AM - 3:00PM	**KDD 2019 Registration** – Tikahtnu Foyer- Level 3, Dena'ina
8:00AM - 9:30AM	**Keynote Address: Professor Cynthia Rudin, Duke University: Do Simpler Models Exist and How Can We Find Them? (see page 19 for full details)** – Tukatnu Ballroom- Level 3, Dena'ina
9:30AM - 5:00PM	**Sponsor Interview Room** – Kodiak Boardroom- Marriott Hotel
9:30AM - 1:30PM	**KDD Exhibit Hall** – Idlughet Hall- Street Level, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: From Shallow to Deep Language Representations: Pre-Training, Fine-Tuning and Beyond** – Kahtnu- Level 3, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: From Graph to Knowledge Graph: Mining Large Scale Heterogeneous Networks Using SPARK** – Tubughnenq 3- Level 3, Dena'ina
9:30AM - 12:30PM	**Hands-On Tutorial: Efficient Data Collection via Crowdsourcing: Aggregation, Incremental Relabeling, and Pricing** – Tubughnenq 5- Level 3, Dena'ina
9:30AM - 10:00AM	**KDD Coffee Break** – Lobby Foyer- Dena'ina
10:00AM - 12:00PM	**Applied Data Science Invited Talks (see page 23 for full schedule)** – Cook- Street Level, Egan
10:00AM - 12:00PM	**Research Track Oral Presentations: RT16 Machine Learning Themes I** – Summit 2- Ground Level, Egan
10:00AM - 12:00PM	**Research Track Oral Presentations: RT17 Interpretability** – Summit 3- Ground Level, Egan
10:00AM - 12:00PM	**Research Track Oral Presentations: RT18 Recommender Systems II** – Summit 4- Ground Level, Egan
10:00AM – 12:00PM	**Applied Data Science Oral Presentations: ADS7 Entity Extraction, Linking, and Search** – Summit 5/6- Ground Level, Egan
10:00AM – 12:00PM	**Applied Data Science Oral Presentations: ADS8 Sensor and Consumer Services** – Summit 1- Ground Level, Egan
12:00PM - 1:30PM	**KDD Lunch** – Idlughet Hall- Street Level, Dena'ina
1:00PM – 4:00PM	**Hands-On Tutorial: From Shallow to Deep Language Representations: Pre-Training, Fine-Tuning and Beyond** - Kahtnu- Level 3, Dena'ina
1:00PM – 4:00PM	**Hands-On Tutorial: From Graph to Knowledge Graph: Mining Large Scale Heterogeneous Networks Using SPARK** – Tubughnenq 3- Level 3, Dena'ina
1:00PM – 4:00PM	**Hands-On Tutorial: Building production-ready recommendation systems at scale** – Tubughnenq 5- Level 3, Dena'ina
1:30PM - 3:30PM	**Applied Data Science Invited Talks (see page 23 for full schedule)** – Cook- Street Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT19 Machine Learning Themes II** – Summit 2- Ground Level, Egan
1:30PM - 3:30PM	**Research Track Oral Presentations: RT20 Online and Incremental Algorithms** – Summit 3- Ground Level, Egan
1:30PM - 3:30PM	**Applied Data Science Oral Presentations: ADS9 E-Commerce and Advertising** – Summit 1- Ground Level, Egan
3:30PM – 4:00PM	**KDD Coffee Break** – Lobby Foyers- Dena'ina
4:00PM - 5:00PM	**KDD 2019 Closing Session** – Tikahtnu Ballroom-Level 3, Dena'ina

GENERAL CHAIRS'

WELCOME MESSAGE

Our society as we know it is going through a fundamental digital transformation. Fueling this transformation is the insights we gain from data in service of one and in service of all. We are holding the future in our hands and therefore share both the excitement and the responsibility of making our innovations useful and trustworthy for each other. Over the last quarter of a century, now in the twenty fifth year since its inception, our community has met each year at the annual conference on Knowledge Discovery and Data Mining – KDD. From humble beginnings of a symposium in 1994, we have become a major movement for innovation in every field where computing plays a role today. Thousands of practical applications fueled by peer-reviewed research that is presented at KDD inform and touch the lives of millions of global citizens. It is time to pause, reflect and celebrate the movement that has grown from a few dozen researchers to a community of tens of thousands of industry and academic participants. It is with this spirit of reflection and celebration that we welcome you to join us at the frontiers of innovation and human spirit in the city of Anchorage, Alaska for KDD 2019.

With such great success comes greater responsibility to ensure we focus on making the world safer, better and more just. Data Mining as a term has largely morphed into data science, and is often used interchangeably with machine learning. The considerations for designing algorithms that learn from larger and more diverse datasets are very different than designing algorithms for small datasets that were a norm just a few decades ago. The importance of end-to-end solutions and role that data science tools play in enabling them is well recognized. While it is obvious to many of us that specific algorithms are but tools, we must resist the temptation to hype them as generalizable intelligence. There is a need for careful debate around making results of the algorithms fair and transparent. We thus need to discuss the impacts of our work in light of its ethics, fairness, and explainability of our tools. The agenda of the conference is thus two fold; focusing on first the richness of algorithmic improvements and second, the impact of the field across our environment, our health, our social interactions and our desire for automation of every day processes from farming to driving to marketing and even as fundamental a task as speaking and listening.

This twenty-fifth meeting of KDD has many firsts; a record number of submissions (400+ more than the prior record), a double blind review process for research track, increased focus on diversity and inclusion through the organization and programmatic elements, a record number of workshops in the program to enable small group interactions, the addition of Earth Day as a special theme day focusing on novel big Earth data, associated Earth data mining challenges (e.g., Gerrymandering risks for data mining results) as well as opportunities (e.g., environmental sustainability, conservation, climate change as well as smart and connected communities), a comprehensive offering of hands-on tutorials from most major machine learning tool vendors, evolving the KDD Cup into a multi-theme contest with significant prize money, and ensuring that we can provide childcare for participants, to name a few.

The first two days of the conference include tutorials by experts in their field and workshops and theme days, followed by three rigorous yet enjoyable days of research papers, applied data science papers and over 20 applied invited speakers sharing views from the trenches of industrial applications. We have also added a keynote plenary panel where the community can come and discuss the why, where and how of democratizing data science and its societal implications. To ensure that the conversation continues long into the night, this panel is followed by the KDD 25th Anniversary Celebration. There are two evening receptions with poster sessions as well to allow attendees to network with each other and discuss technical details of papers in small group settings. All this is made possible with the generous support of our sponsors who have, to date, pledged over USD 1M in sponsorships! We thank them.

KDD is completely volunteer run and managed. Behind the scenes, ensuring this success is a team of committed organizers who share the passion for data science and tirelessly give back to make the conference a huge success. With the spirit of the wild that Alaska is known for and is a true inspiration for all researchers, academics, industry practitioners, and entrepreneurs in data science, machine learning and AI, we welcome you to participate and share your learnings at KDD 2019.

Ankur Teredesai & Vipin Kumar

PROGRAM CHAIRS'

WELCOME MESSAGE

Welcome to the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. The record number of papers submitted and accepted in this year's conference highlights the breadth, depth, and impact of the research and applied data science practices on our vibrant fields in this exciting time in history. It is with great pleasure and joyful anticipation that we present this proceedings to you. We have given our uttermost best effort to ensure that these papers meet high standards of quality in terms of presentation, citations, and experimental methodology. We sincerely hope that you will enjoy this year's program and that you will find the papers insightful, inspiring, and energizing. We expect them to contribute to the foundation and tool chest that you can draw value from for your future contributions to the fields of knowledge discovery and data mining.

This year's KDD continues the trend of record high number of submissions, a total of 1808 submissions, out of which 321 were accepted for publication in these proceedings; thus reaching an acceptance rate of 17.8%. The submissions came from a total 7966 authors from over 1200 organizations across 58 countries.

One of the changes in this year's Research Track is the introduction of a double-blind review process. Double-blind reviewing has been adopted by many conferences in our field, and SIGKDD's Executive Committee (EC) along with the KDD 2019's General Chairs and PC Chairs felt that the time has come for KDD to also embrace this review process. Double-blind reviewing is a tool that attempts to eliminate implicit biases that all of us possess and can influence our reviews in imperceptible ways. To a large extent, the double-blind review process went smoothly. We hope that this is a first step towards improving the reviewing process and that this initial attempt will stimulate further discussions and allow SIGKDD to design a permanent policy for double-blind reviewing.

The Applied Data Science Track witnessed a higher growth rate in terms of submissions, with approximtely 40% more papers submitted to the track this year relative to last year. In addition, for the first time in KDD, we introduced mandatory training sessions for all Program Committee members, prior to the start of the review process. This allowed us to share best practices about reviewing applied work, emphasizing the key role of the track in advancing the understanding of issues related to deploying/leveraging these technologies in the real world. We were excited to see that over 90% of the reviewers went through the training. We also introduced a step in the submission for authors to declare their paper topics. This provided additional insight about the diverse topics that we are working on and helped us identify reviewers during the reviewing process. Finally, like the research track, given the importance of reproducibility in data mining work, this year we emphasized reproducibility and allowed authors to use up to two extra pages to include information that would help readers reproduce the work.

We would like to acknowledge the tremendous work of the Senior Program Committee members (85 for the Research Track, 32 for the Applied Data Science Track), the Program Committee members (627 for the Research Track, 222 for the Applied Data Science Track), and more than 800 external reviewers. Together they wrote well over 6500 reviews, and held multiple rounds of discussions for most of the submissions. We are especially indebted to the Senior Program Committee members for their guidance to PC members and valuable input to us.

We would also like to thank the general chairs, Vipin Kumar and Ankur Teredesai, and the rest of the organizing committee. Thank you for the great team work and for a joyful time together.

Lastly but most importantly, we are grateful to all authors who made the submissions, and attendees at KDD. Your work and your presence are part of what made KDD one of the best conferences we all love. We look forward to meeting you in person at the conference and in future venues.

Ying Li and Rómer Rosales

KDD'19 Applied Data Science Track Chairs

Evimaria Terzi and George Karypis

KDD'19 Research Track Chairs

PROGRAM HIGHLIGHTS

Opening Ceremony and Award Presentations

Opening Welcome Address by General Co-Chairs

KDD 2019 Award Presentations:

Best Paper awards in Research and Applied Data Science, KDD Cup Winners,

SIGKDD Awards:

Innovation Award and Talk – **Charu Aggarwal**, Service Award – **Balaji Krishnapuram**, Test-of-Time Award – **Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutos, Jeanne VanBriesen, Natalie Glance**. Cost-Effective Outbreak Detection in Networks, Best Dissertation Award – **Tim Althof**, Stanford, advised by **Jure Leskovec**

KDD 2019 Keynote Talks

Peter Lee. Corporate Vice President, Microsoft. The Unreasonable Effectiveness, and Difficulty, of Data in Healthcare. Introduced by General Co-Chair **Ankur Teredesai**, Co-Founder CTO, KenSci & Professor of Computer Science and Systems, University of Washington.

Cynthia Rudin. Professor of Computer Science, Electrical and Computer Engineering, and Statistical Science (Duke University). Do Simpler Models Exist and How Can We Find Them? Introduced by General Co-Chair **Vipin Kumar**, Professor of Computer Science, University of Minnesota.

KDD 25th Anniversary Plenary Keynote Panel

Title: **Why, What and How we Democratize Data Science?**

Panelists: **Tina Eliassi-Rad**, Northeastern University, **Yu Zheng**, JD.com, **Eric Sears**, MacArthur Foundation, **Rediet Abebe**, Cornell University & **Anil Jain**, Michigan State University. Moderated by **Nitesh Chawla**, University of Notre Dame.

Presentations by Authors of Peer-Reviewed Research and Applied Data Science Track Papers

- 174 Research Track Papers. Acceptance Rate is 15%
- 148 Applied Data Science Track Papers. Acceptance Rate is 15%

Industry Invited Talks

21 Information-Filled Talks sharing the view in the trenches, by select Data Science Industry Thought

Lecture-Style Tutorials

29 In-Depth tutorials covering a range of basic and advanced topics in data science research

Hands-On Tutorials

15 hands-on labs based tutorials featuring the latest in machine learning tools and platforms

KDD Workshops

34 Workshops in various data science sub-fields with carefully selected papers that describe latest advances including works in progress and early results in these areas.

Deep Learning Day

4 Keynote talks followed by a plenary panel: **Dawn Song** (UC Berkeley), **Jie Tang** (Tsinghua), **Ruslan Salakhutdinov** (CMU & Apple), **Jure Leskovec** (Stanford & Pinterest)

4 Workshops:

- Reinforcement learning
- High-dimensional sparse data
- Graphs
- Education

Posters Session featuring a selection of papers (jointly with the main KDD Poster Session)

Earth Day

2 Keynotes and plenary panels:

- Emerging Cloud-based Big Earth Data Initiatives – **Dr. Rama Nemani**, NASA Earth Exchange
- Earth Data Mining Challenge Landscape – **Prof. Harvey Miller**, UCGIS Call for Action

3 Workshops:

- Fragile Earth: Theory Guided Data Science to Enhance Scientific Discovery
- Data Mining and AI for Conservation
- The 8th SIGKDD International Workshop on Urban Computing

Posters Session featuring a selection of papers (jointly with the main KDD Poster Session)

Health Day

Opening Keynote: Dr. Eric Eskioglu, Novant Health

Plenary Panel Discussion: Healthcare Beyond the 4 walls: An AI opportunity

Dr. Anthony Chang, Children's Hospital Orange County, **Dr. Diana Bottari**, Advocate Children's Hospital, and other notable expert panelists. Moderated by **Dr. Carly Eckert**, KenSci.

2 Keynote Talks:

- **Srinivas Aluru**, Georgia Institute of Technology
- **Tae Hyun Hwang**, Cleveland Clinic Lerner College of Medicine

3 Workshops:

- DS Health: Relevance and Importance of Knowledge-drive AI
- bioKDD: Workshop on Data Mining in Bioinformatics
- epiDAMIK: Epidemiology meets Data Mining and Knowledge Discovery

Posters Session featuring a selection of papers (jointly with the main KDD Poster Session)

KDD Cup Day

3 Contests with $100K in prizes

Track 1: Regular Machine Learning Competition Track (Regular ML Track)

Track 2: Automated Machine Learning Competition Track (AutoML Track)

Track 3: "Research for Humanity" Reinforcement Learning Competition Track (Humanity RL Track)

Contest winning strategy and approach presentations by top 3 teams in each track

Project Showcase

Featuring global large collaborative research programs.

2 Keynotes Talks:

- **Dennis Pamlin**, Research Institutes of Sweden
- **Stefano Bertolo**, European Commission

11 Project Presentations with posters and live demonstrations by project team members

Social Impact Workshop

2 Invited Keynotes:

- **Karen Matthys**, Stanford University & Women in Data Science movement
- **Anissa Tanweer**, University of Washington

8 Global social impact project presentations

LECTURE-STYLE

TUTORIAL DAY

Chairs: Danai Koutra (University of Michigan, Ann Arbor) and Leman Akoglu (Carnegie Mellon University)

Sunday August 4th, 2019

***Locations and times provided in the Agenda at a Glance ***

T1: Learning From Networks: Algorithms, Theory, & Applications. Xiao Huang (Texas A&M), Peng Cui (Tsinghua), Yuxiao Dong (Microsoft), Jundong Li (Arizona State/University of Virginia), Huan Liu (Arizona State), Jian Pei (Simon Fraser University), Le Song (Georgia Tech), Jie Tang (Tsinghua), Fei Wang (Cornell University), Hongxia Yang (Alibaba), Wenwu Zhu (Tsinghua)

T2: Data Integration and Machine Learning: A Natural Synergy. Xin Luna Dong (Amazon), Theodoros Rekatsinas (University of Wisconsin-Madison)

T3: Mining and model understanding on medical data. Myra Spiliopoulou (University Magdeburg), Panos Papapetrou (Stockholm University)

T4: Deep Reinforcement Learning with Applications in Transportation. Zhiwei (Tony) Qin (Didi Chuxing), Jian Tang (Didi Chuxing & Syracuse University), Jieping Ye (Didi Chuxing & University of Michigan, Ann Arbor)

T5: Gold Panning from the Mess: Rare Category Exploration, Exposition, Representation and Interpretation. Dawei Zhou (Arizona State University), Jingrui He (Arizona State University)

T6: Social User Interest Mining: Methods and Applications. Fattane Zarrinkalam (Ryerson University), Hossein Fani (University of New Brunswick), Ebrahim Bagheri (Ryerson University)

T7: Fairness-Aware Machine Learning: Practical Challenges and Lessons Learned. Sarah Bird (Microsoft, USA), Ben Hutchinson (Google, USA), Krishnaram Kenthapadi (LinkedIn, USA), Emre Kıcıman (Microsoft, USA), Margaret Mitchell (Google, USA)

T8: Modern MDL meets Data Mining -- Insights, Theory, and Practice. Jilles Vreeken (Helmholtz Center for Information Security), Kenji Yamanishi (U. Tokyo)

T9: Adaptive Influence Maximization. Bogdan Cautis (U Paris-Sud, France), Silviu Maniu (U Paris-Sud, France), Nikolaos Tziortziotis (Tradelab and U Paris-Sud, France)

T10: Hypothesis Testing and Statistically-sound Pattern Mining. Leonardo Pellegrina (University of Padova, Italy), Matteo Riondato (Amherst College, USA), Fabio Vandin (University of Padova, Italy)

T11: Fake News Research: Theories, Detection Strategies, and Open Problems. Reza Zafarani (Syracuse University), Xinyi Zhou (Syracuse University), Kai Shu (Arizona State University), Huan Liu (Arizona State University)

T12: Are You My Neighbor? Bringing Order to Neighbor Computing Problems. David C. Anastasiu (San Jose State University), Huzefa Rangwala (George Mason University), Andrea Tagarelli (University of Calabria, Italy)

T13: Mining temporal networks. Aristides Gionis (Aalto University), Polina Rozenshtein (Nordea Data Science Lab)

T14: Modeling and Applications for Temporal Point Processes. Junchi Yan (SJTU), Hongteng Xu (Duke & Infinia ML, Inc), Liangda Li (Yahoo Research)

T15: Challenges, Best Practices and Pitfalls in Evaluating Results of Online Controlled Experiments. Xiaolin Shi (Snap Inc), Pavel Dmitriev (Outreach), Somit Gupta (Microsoft), Xin Fu (Facebook)

T16: Foundations of large-scale sequential experimentation. Aaditya Ramdas (Carnegie Mellon University)

T17: Constructing and Mining Heterogeneous Information Networks from Massive Text. Jingbo Shang (UIUC), Jiaming Shen (UIUC), Liyuan Liu (UIUC), Jiawei Han (UIUC)

T18: Deep Bayesian Mining, Learning and Understanding. Jen-Tzung Chien (National Chiao Tung University, Taiwan)

T19: Incompleteness in Networks: Biases, Skewed Results, and Some Solutions. Tina Eliassi-Rad (Northeastern University), Rajmonda Caceres (MIT Lincoln Laboratory), Timothy LaRock (Northeastern University)

T20: Optimize the Wisdom of the Crowd: Inference, Learning, and Teaching. Yao Zhou (Arizona State University), Fenglong Ma (PSU), Jing Gao (U Buffalo), Jingrui He (Arizona State University)

T21: Interpretable Knowledge Discovery Reinforced by Visual Methods. Boris Kovalerchuk (CWU)

T22: Explainable AI in Industry. Krishna Gade (Fiddler Labs, USA), Sahin Cem Geyik (LinkedIn, USA), Krishnaram Kenthapadi (LinkedIn, USA), Varun Mithal (LinkedIn, USA)

T23: Advances in Cost-sensitive Multiclass and Multilabel Classification. Hsuan-Tien Lin (National Taiwan University)

T24: Recent Progress in Zeroth Order Optimization and Its Applications to Adversarial Robustness in Data Mining and Machine Learning. Pin-Yu Chen (IBM Research), Sijia Liu (IBM Research)

T25: Forecasting Big Time Series: Theory and Practice. Christos Faloutsos (CMU and Amazon), Valentin Flunkert (AWS AI Labs), Jan Gasthaus (AWS AI Labs), Tim Januschowski (AWS AI Labs), Yuyang (Bernie) Wang (AWS AI Labs)

T26: Deep Natural Language Processing for Search and Recommender Systems. Weiwei Guo (LinkedIn), Huiji Gao (LinkedIn), Jun Shi (LinkedIn), Bo Long (LinkedIn), Liang Zhang (LinkedIn), Bee-Chung Chen (LinkedIn), Deepak Agarwal (LinkedIn)

T27: **Spatio-temporal event forecasting and precursor identification.** Yue Ning (Stevens Institute of Technology), Liang Zhao (George Mason University), Feng Chen (SUNY, Albany), Chang-Tien Lu (Virginia Tech), Huzefa Rangwala (George Mason University)

T28: Statistical Mechanics Methods for Discovering Knowledge from Production-Scale Neural Networks. Charles Martin (Calculation Consulting), Michael Mahoney (UC Berkeley)

T29: Data Mining Methods for Drug Discovery and Development. Cao (Danica) Xiao (IQVIA), Jimeng Sun (Georgia Tech)

WORKSHOP DAY

Chairs: Anuj Karpatne (Virginia Tech University) and Jing Gao (University of Buffalo, SUNY)

Monday August 5, 2019

***Locations provided in the Agenda at a Glance ***

8:00AM - 5:00PM (Full Day)

- W1: The 5th Workshop on Mining and Learning from Time Series
- W2: AdKDD 2019 Workshop
- W3: 15th International Workshop on Mining and Learning with Graphs (MLG)

8:00AM - 12:00PM (Half Day)

- W4: [Earth Day]: Fragile Earth: Theory Guided Data Science to Enhance Scientific Discovery
- W5: [Earth Day]: Data Mining and AI for Conservation
- W6: [Earth Day]: The 8th SIGKDD International Workshop on Urban Computing
- W7: The 8th International Workshop on Parallel and Distributed Computing for Large-Scale Machine Learning and Big Data Analytics
- W8: The 1st International Workshop on Artificial Intelligence of Things
- W9: 2nd Workshop on Offline and Online Evaluation of Interactive Systems
- W10: KDD 2019 Workshop on Causal Discovery (CD2019)
- W11: The First International Workshop on Smart Data for Blockchain and Distributed Ledger (SDBD'19)
- W12: Tensor Methods for Emerging Data Science Challenges
- W13: 2nd KDD Workshop on Anomaly Detection in Finance
- W14: Explainable AI/ML (XAI) for Accountability, Fairness, and Transparency
- W15: AI for fashion : The fourth international workshop on fashion and KDD
- W16: Data Collection, Curation, and Labeling for Mining and Learning
- W17: Adversarial Learning Methods for Machine Learning and Data Mining
- W18: 2019 International Workshop on Talent and Management Computing (TMC'2019)

1:00PM - 5:00PM (Half Day)

- W19: [Deep Learning Day]: Deep Learning on Graphs: Methods and Applications (DLG'19)
- W20: [Deep Learning Day]: Deep Reinforcement Learning for Knowledge Discovery
- W21: [Deep Learning Day]: 1st International Workshop on Deep Learning Practice for High-Dimensional Sparse Data with KDD 2019
- W22: [Deep Learning Day]: Deep Learning for Education (DL4Ed)
- W23: [Health Day]: [DsHealth 2019] 2019 KDD workshop on Applied data science in Healthcare: bridging the gap between data and knowledge
- W24: [Health Day]: epiDAMIK: Epidemiology meets Data Mining and Knowledge discovery
- W25: [Health Day]: 18th International Workshop on Data Mining in Bioinformatics (BIOKDD'19)
- W26: Developing Standards on Definitions of Analytics Roles, Skill-sets and Career Paths in the Data Science Industry
- W27: The First International Workshop on Intelligent Information Feed
- W28: The Third International Workshop on Automation in Machine Learning
- W29: BigMine-19: 8th International Workshop on Big Data, IoT Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications
- W30: Learning and Mining for Cyber Security (LEMINCS)
- W31: From Data Integration to Knowledge Graphs: Challenges and Experiences
- W32: Data-driven Intelligent Transportation (DIT'19)
- W33: Truth Discovery and Fact Checking: Theory and Practice
- W34: Eighth International Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM'19)

THEME:

EARTH DAY

Monday, August 5, 2019

Locations provided in the Agenda at a Glance

Earth Day will bring together thought leaders in academia, industry and government to explore this area and discuss opportunities to overcome the challenges that Earth faces today.

Chairs: Shashi Shekhar, University of Minnesota – Twin Cities; James Hodson, AI for Good Foundation; Lucas Joppa, Microsoft Research

8:00AM - 12:00PM EARTH-RELATED WORKSHOPS

- Fragile Earth: Theory Guided Data Science to Enhance Scientific Discovery
- Data Mining and AI for Conservation
- Urban Computing

12:00 PM - 1:00 PM Lunch Provided in Exhibit Hall

1:30PM - 1:45PM WELCOME, WHY EARTH DAY? WHY NOW?

1:45P - 3:00PM SESSION 1: BIG EARTH DATA SETS & RECENT INITIATIVES

Background: Earth data (e.g., remote sensing imagery, GPS time service, location traces) has already transformed our lives by improving monitoring of global weather and agriculture for early warning of hurricanes and inclement weather as well as food shortage risks due to crop stresses or failures. Further, with 2 billion [2] receivers in use for location and time services, the GPS has become a critical infrastructure for the world economy for use cases ranging from precision agriculture to navigation to ride sharing to smart cities. Moreover, billions of weather forecasts are used every day worldwide leveraging data from Earth Observation satellites. These success stories are only a beginning and many transformative opportunities lie ahead. For example, the 2011 Mckinsey Big Data report [4] estimated that location trace data will generate about $600 billion annually by 2020. In addition, 2019 U.S. national academy report projects $1.6 trillion in savings for energy generation and use from earth data by 2035 [1]. Furthermore, government and industry have recently started major initiatives such as NASA Earth Exchange, Amazon's Earth on AWS, Google Earth Engine, Microsoft's AI for Earth, and NSF Navigating the New Arctic for meeting grand challenges facing our changing planet such as conservation, climate change, and environmental sustainability. This session will explore the tremendous value of Earth data for civil society, prosperity and good governance via a keynote and a panel discussion.

Keynote (30 min): Dr. Ramakrishna Nemani, NASA Earth Exchange

Panel (45 min): Societal value of Earth Data (Chair: James Hodson, AI for Good Foundation)

- Jennifer Marsman, Microsoft AI for Earth
- Dennis Pamlin, RISE Research Institutes of Sweden
- Dr. Robert Stewart, OakRidge National Laboratory, USDOE
- Rob Bochenek, NOAA Integrated Ocean Observing System & Axiom Data Science

4:00PM - 5:15PM SESSION 2 : EARTH DATA MINING: CHALLENGES AND OPPORTUNITIES

Background: Data mining methods have found success in analyzing many complicated systems, such as e-commerce, and use cases explored in the Earth Day aligned SIGKDD workshops. However, many questions remain open due to unique Earth data challenges such as spatio-temporal auto-correlation, heterogeneity, scale-dependence, measurement errors, modifiable areal unit problem, etc [2,3,5]. For example, a recent Geo-Physical letter paper [1] noted that "failure to account for dependence between [Physical] models, variables, locations and seasons yield misleading results". Additional challenges are noted in recent community papers from the NSF IS-GEO Research Coordination Network [2] and University Consortium for Geographic Information Science [3]. For example, Gerrymandering court debates [6] also raise transparency concerns for the risk of altering statistical results by changing the choice of spatial partitions. This session explores these challenges and opportunities via a keynote and a panel discussion.

Keynote (30 min): Prof. Harvey Miller, Director, Center for Urban and Regional Analysis, OSU

Panel (45 min): Earth Data Mining: Challenges and Opportunities

(Chair: Prof. Shashi Shekhar, McKnight Distinguished University Professor, University of Minnesota)

- Prof. Dan Griffith, AAAS Fellow, University of Texas, Dallas
- Prof. Anuj Karpatne, Virginia Tech
- Prof. Tanya Berger-Wolf, University of Illinois at Chicago
- Dr. Srinivas (Sai) Ravela, Massachusetts Institute of Technology

THEME:

DEEP LEARNING DAY

Monday, August 5, 2019

Locations provided in the Agenda at a Glance

Deep Learning Day is specifically dedicated to deep learning's impact on data science. The goal is to provide an overview of recent developments in deep learning, particularly for data mining.

Chairs: Yuxiao Dong, Microsoft Research; Vagelis Papalexakis, University of California Riverside

8:00AM - 12:00PM Deep Learning Day Plenary Keynotes - Egan Center

- 8:00 - 8:45 Dawn Song (UC Berkeley)
- 8:45 - 9:30 Jie Tang (Tsinghua University)
- 9:30 - 10:00 Coffee Break
- 10:00 - 10:45 Ruslan Salakhutdinov (CMU & Apple)
- 10:45 - 11:30 Jure Leskovec (Stanford & Pinterest)

11:30AM - 12:00PM Deep Learning Day Panel with Dawn Song, Jie Tang, Ruslan Salakhutdinov, and Jure Leskovec

12:00PM - 1:00PM Lunch Provided in Exhibit Hall

1:00PM - 5:00PM Deep Learning Day Workshops

- Workshop on Deep Reinforcement Learning for Knowledge Discovery
- Workshop on Deep Learning Practice for High-Dimensional Sparse Data
- Workshop on Deep Learning on Graphs: Methods & Applications
- Workshop on Deep Learning for Education

THEME:

HEALTH DAY

Monday, August 5, 2019

Locations provided in the Agenda at a Glance

Health Day is dedicated to bringing together domain and machine learning experts to discuss challenges and trends in the healthcare industry, as well as techniques and methodologies machine learning community.

Chairs: Faisal Farooq, IBM; Carly Eckert, Kensci; Chandan Reddy, Virginia Tech

8:00AM - 12:00PM Health Day Plenary Session

- 8:00 - 8:30 Opening Remarks
- 8:30 - 9:30 **Plenary Keynote:** Dr. Eric Eskioglu, Novant Health
- 9:30 - 10:00 Coffee Break
- 10:00-10:30 **Keynote Talk:** Tae Hyun Hwang, PI Hwang Lab, Cleveland Clinic Lerner College of Medicine
- 10:30 - 11:00 **Keynote Talk:** Srinivas Aluru, Professor, Georgia Institute of Technology
- 11:00 - 12:00 **Plenary Panel:** Healthcare Beyond the Four Walls: an AI opportunity.

  The future of healthcare will exist beyond the walls of hospitals and doctors' offices. Telemedicine, precision medicine, consumer health, and digital health are all advances in healthcare designed to extend beyond the traditional clinical setting. Innovative approaches are needed to deliver care in unorthodox settings and in novel arrangements. What are the opportunities for AI in these situations? How can AI serve to "meet patients where they are" and help people find and use resources to make healthcare work for their advantage? We will be joined by a stellar panel of doctors and data scientists with experience in these fields. These are practitioners who are experienced in the challenges, successes, and the pitfalls of applying AI in the community or home setting.

  We expect a lively panel with a diversity of perspectives to address these issues and others such as:

  - How can digital health and precision medicine be used in underserved populations?
  - What are the frontiers for behavioral and mental health care?
  - Are there opportunities for AI/ML to augment healthcare in austere environments?
  - What are the realities of the digital divide and how does it affect your work?
  - What are the unexpected barriers of engagement with patients (and their families)?
  - What additional considerations are needed when protected populations are involved

12:00PM - 1:00PM Lunch Provided in Exhibit Hall

1:00PM - 5:00PM Health Day Workshops

- [DsHealth 2019] 2019 KDD workshop on Applied data science in Healthcare: bridging the gap between data and knowledge
- [epiDAMIK 2019] Epidemiology meets Data Mining and Knowledge discovery
- [bioKDD 2019] 18th International Workshop on Data Mining in Bioinformatics (BIOKDD'19)

KDD CUP DAY

Tuesday, August 6, 2019

10:00AM-5:00PM, Arteaga Room, Street Level, Egan Center

KDD Cup is the annual data mining and knowledge discovery competition organized by ACM Special Interest Group on Knowledge Discovery and Data Mining, the leading professional organization of data miners.

Chairs: Taposh Roy, Kaiser Permanente; Iryna Skrypnyk, Pfizer; Wenjun Zhou, University of Tennessee Knoxville

10:00 - 10:30 Introduction, KDD Cup Chairs

10:30 - 11:25 Winner Announcements

Introduction (Iryna Skrypnyk)

Baidu: Regular ML competition (Awards #1, #2 and #3)

Introduction (Taposh Roy)

4Paradigm : AutoML Competition (Awards #1, #2 and #3)

Introduction (Wenjun Zhou)

IBM Africa & Hexagon-ML : Reinforcement Learning Competition (Awards #1, #2 and #3)

11:30 - 11:45 KDD Cup - Innovation Award

Awarded by RL Guru : Dr. Balaraman Ravindran, IIT Madras

KDD Cup Innovation award is for getting new ideas into data science and machine learning competition space. This year, the RL competition introduced and attracted participation from more than 300 individuals across the world.

11:45 - 1:00 Lunch is Provided in Exhibit Hall

1:00 - 3:15 Spotlight Talks

1:00 - 1:45 How did I win the AutoML competition?

Introduction: Iryna Skrypnyk

Speaker: AutoML competition winners and 4Paradigm

1:45 - 2:30 How did I win the ML competition?

Introduction: Wenjun Zhou

Speaker: Regular ML competition winners and Baidu

2:30 - 3:15 How did I win the Humanity RL competition?

Introduction: Taposh Roy

Speaker: RL competition winners and IBM Africa

3:30 - 4:00 Panel 1: How should companies use competition platforms "internal and/or external"?

4:15 - 5:00 Panel 2: How will AutoML change the future of data science?

PROJECT SHOWCASE PROGRAM

Wednesday, August 7, 2019

8:30AM-3:30PM, Idlughet Hall 3, Street Level, Dena'ina Center

The Project Showcase Track offers a full day focused exclusively on innovative KDD-relevant projects from national and regional funding programs, as well as corporate, start-up, and nonprofit channels.

Chairs: Dunja Mladenic, Jožef Stefan Institute; James Hodson, Jožef Stefan Institute

Agenda

- 8:30 - 9:10 Keynote: Challenges of Data Mining Projects
- 9:10 - 9:30 Data & Frameworks (2 projects, 10 mins each)
- 9:30 - 10:00 Coffee break
- 10:00 - 11:10 Data & Society (6 projects, 10 mins each)
- 11:10 - 12:00 Demos I (8 projects + 2 posters)
- 12:00 - 1:30 Lunch break
- 1:30 - 2:10 Keynote: Dennis Pamlin, "AI, Big Data, and the UN Sustainable Development Agenda"
- 2:10 - 2:40 Data for Earth Sensing (3 projects, 10 mins each)
- 2:40 - 3:30 Demos II (the presented 3 projects + 5 posters)

KEYNOTES

Peter Lee

Microsoft Healthcare
Redmond, WA, USA
@peteratmsr

The Unreasonable Effectiveness, and Difficulty, of Data in Healthcare

Tuesday, August 6th 8:00AM - 9:30AM

ABSTRACT

Data and data analysis are widely assumed to be the key part of the solution to healthcare systems' problems. Indeed, there are countless ways in which data can be converted into better medical diagnostic tools, more effective therapeutics, and improved productivity for clinicians. But while there is clearly great potential, some big challenges remain to make this all a reality, including making access to health data easier, addressing privacy and ethics concerns, and ensuring the clinical safety of "learning" systems. This talk illustrates what is possible in healthcare technology, and details key challenges that currently prevent this from becoming a reality.

BIOGRAPHY

Dr. Peter Lee is Corporate Vice President, Microsoft Healthcare. He leads an organization that works on technologies for better and more efficient healthcare, with a special focus on artificial intelligence and cloud computing. Dr. Lee has extensive experience in managing the process of going from basic research to commercial impact. Past illustrative examples include the deep neural networks for simultaneous language translation in Skype, next-generation IoT technologies, and innovative silicon and postsilicon computer architectures for Microsoft's cloud. He also has a history of advancing more "out of the box" technical efforts, such as experimental under-sea datacenters, augmented-reality experiences for HoloLens and VR devices, digital storage in DNA, and social chatbots such as XiaoIce and Tay. Previously, as an Office Director at DARPA, he led efforts that created operational capabilities in advanced machine learning, crowdsourcing, and bigdata analytics, such as the DARPA Network Challenge and Nexus 7. He was formerly the Head of Carnegie Mellon University's computer science department. As a thought leader, he has spoken and written widely on technology trends and policies, spanning the fields of computing technology, healthcare, and innovation ecosystem. He is a member of the Boards of Directors of the Allen Institute for Artificial Intelligence and the Kaiser Permanente School of Medicine. He served on President's Commission on Enhancing National Cybersecurity. He has led studies for the National Academies on the impact of federal research investments on economic growth and testified before the US House Science and Technology Committee and the US Senate Commerce Committee.

Cynthia Rudin

Duke University
Durham, NC
cynthia@cs.duke.edu

Do Simpler Models Exist and How Can We Find Them?

Thursday, August 8th 8:00AM - 9:00AM

ABSTRACT

While the trend in machine learning has tended towards more complex hypothesis spaces, it is not clear that this extra complexity is always necessary or helpful for many domains. In particular, models and their predictions are often made easier to understand by adding interpretability constraints. These constraints shrink the hypothesis space; that is, they make the model simpler. Statistical learning theory suggests that generalization may be improved as a result as well. However, adding extra constraints can make optimization (exponentially) harder. For instance it is much easier in practice to create an accurate neural network than an accurate and sparse decision tree. We address the following question: Can we show that a simple-but-accurate machine learning model might exist for our problem, before actually finding it? If the answer is promising, it would then be worthwhile to solve the harder constrained optimization problem to find such a model. In this talk, I present an easy calculation to check for the possibility of a simpler model. This calculation indicates that simpler-but-accurate models do exist in practice more often than you might think. I then briefly overview several new methods for interpretable machine learning. These methods are for (i) sparse optimal decision trees, (ii) sparse linear integer models (also called medical scoring systems), and (iii) interpretable case-based reasoning in deep neural networks for computer vision.

BIOGRAPHY

Cynthia Rudin is a professor of computer science, electrical and computer engineering, and statistical science at Duke University, and directs the Prediction Analysis Lab, whose main focus is in interpretable machine learning. 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 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. She 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, and AAAI. She has served on three 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 Electric Grid. She is a fellow of the American Statistical Association and a fellow of the Institute of Mathematical Statistics. She will be the Thomas Langford Lecturer at Duke University during the 2019-2020 academic year.

KEYNOTE PANEL

Chairs: Chid Apte (IBM Research), Geoff Webb (Monash University) and Nitesh Chawla (University of Notre Dame)

Wednesday, August 7th, 4:00PM – 6:00PM

Tukatnu Ballroom, Level 3, Dena'ina Center

- **Moderator: Nitesh Chawla**, University of Notre Dame
- **Rediet Abebe**, Cornell University
- **Tina Eliasi-Rad**, Northeastern
- **Anil Jain**, Michigan State
- **Eric Sears**, MacArthur Foundation
- **Yu Zheng**, JD.com

Abstract: The beneficiaries of advanced data science remain largely the key stakeholders in various industries. However, as we infuse data science into assisted and automated intelligence services, it will drive a major shift towards data science benefiting the consumers, a democratization of the technology. This panel debates what it would mean to democratize data science, how this might be done and the arguments for and against it. Issues to be covered include:

- What does democratization of data science mean? How might it be achieved?
- What are the tenets of fairness and transparency that must accompany this democratization?
- What (or how will these) applications will be most immediately and profoundly affected by democratization?
  - Driverless cars
  - Healthcare
  - Sustainability
  - Education
  - Environment
  - Security
- What innovations do we anticipate on data sensing and gathering?
- Whose interests would be served by the democratization of data science and whose are served by its absence?
- How do we prepare ourselves for unanticipated risks?
- Do citizens have a right to share in the downstream benefits derived from data they generate?
- Do commercial interests have a right to benefit from data private citizens generate?
- What data ought to be public and what should remain private?
- Will it become a race to data ownership?
- Ethics, trust, and privacy issues
- What is our expectation of deep learning architectures? Is it a computational speed-up and scale challenge or what are the new algorithmic challenges that one may envision? What new learning paradigms await us?
- What will be the data science vs data trade-off? Where will the innovations come from?

KDD 2019

WOMEN'S LUNCH

Chair: Hema Raghavan (LinkedIn)

Wednesday, August 7th, 12:00PM – 1:30PM (Tickets Required)

Arteaga Room, Street Level, Egan Center

KDD will be hosting a panel discussion during the Women's Lunch. The panelists are Rukmini Iyer (Distinguished Engineer, Microsoft), Tina Eliassi-Rad (Associate Professor, Northeastern University), Lillian Carrasquillo (Senior Data Scientist, Spotify), Romer Rosales-Demoral (Senior Director, LinkedIn). The group will discuss a range of topics ranging from an understanding of why environments fostering diversity, inclusion and belonging are important in academia and industry, to what it takes to have a successful career in the field of Data Mining and Artificial Intelligence. Hema Raghavan from LinkedIn will be moderating the panel.

KDD

SOCIAL IMPACT WORKSHOP

Chairs: Vani Mandava (Microsoft Research) and Sarah Stone (EScience Institute)

Wednesday, August 7th, 8:15AM – 11:45AM

8:15 - 8:45 **Karen Matthys** (Stanford University) Invited talk: **Diversity & Social Impact: the WiDS Global Movement**

8:45 - 9:15 **Anissa Tanweer** (University of Washington) Invited talk: **From intentions to impact: Articulating and embracing social complexity in Data Science for Social Good**

9:15 - 9:30 Should You Open-Source Your Model? Ethical questions for open-sourcing Machine Learning models. **Robert Munro** (Lilt, Inc.)

9:30 - 10:00 Coffee Break

10:00 - 10:15 Scikid-Learn: An AI tool for customized education. **Andy Spezzatti** (UC Berkeley), **Mike Lawrence** (UC Berkeley)**, Shine Shan** (UC Berkeley)**, Niema El Bouri** (UC Berkeley)**, Ada Tanyindawn** (UC Berkeley)**, Aljaz Kosmerlj** (AI for Good Foundation)**, James Hodson** (AI for Good Foundation)

10:15 - 10:30 AI-Guided Virtual Lab for Autonomous Vehicle Test: Self-Play Reinforcement Learning Based Two-Player-Game. **Zhaobin Mo**; **Xuan Di** (Columbia Univ)

10:30 - 10:45 Mitigating Demographic Biases in Social Media-Based Recommender Systems. **Rashidul Islam; Kamrun Naher Keya; Shimei Pan; James Foulds** (Univ of Maryland)

10:45 - 11:00 Communicating Machine Learning Results About the Flint Water Crisis to City Residents at Scale. **Jared Webb** (Univ of Michigan); **Eric Schwartz** (University of Michigan); **Jacob Abernethy** (Georgia Tech)

11:00 - 11:15 Their Futures Matter Family Investment Model. **Peter Mulquiney** (Taylor Fry)

11:15 - 11:30 Machine Learning for Humanitarian Data: Tag Prediction using the HXL Standard. **Vinitra Swamy** (Microsoft); **Elisa Chen** (UC Berkeley); **Anish Vankayalapati** (UC Berkeley), **Abhay Aggarwal** (UC Berkeley), **Chloe Liu** (UC Berkeley), **Vani Mandava** (Microsoft), **Simon Johnson** (UN Office of Humanitarian Affairs)

11:30 - 11:45 Mining large-scale news articles for predicting forced migration. **Sadra Abrishamkar** (York University); **Forouq Khonsari** (York University)

KDD 2019

HANDS-ON TUTORIAL PROGRAM

Chairs: Shenghua Bao (Amazon) and Wee Hyong Tok (Microsoft)

Tuesday, August 6th, 2019

Put Deep Learning to Work: A Practical Introduction using Amazon Web Services: Wenming Ye (AWS), Miro Enev (Nvidia)

Democratizing & Accelerating AI through Automated Machine Learning: Parashar Shah (Microsoft), Krishna Anumalasetty (Microsoft)

Concept to Code: Deep Neural Conversational System: Omprakash Sonie (Flipkart), Abir Chakraborty (Flipkart)

Cloud-Based Data Science at the Speed of Thought Using RAPIDS - the Open GPU Data Science Ecosystem on Azure: Brad Rees (NVIDIA), Bartley Richardson (NVIDIA), Tom Drabas (Microsoft), Keith Kraus (NVIDIA), Corey Nolet (NVIDIA), Juan-Arturo Herrera (Microsoft) Kaarthik Sivashanmugam (Microsoft)

Introduction to computer vision and realtime deep learning-based object detection: James G. Shanahan (University of California, Berkeley), Liang Dai (Facebook, University of California Santa Cruz)

Declarative Text Understanding with SystemT: Huaiyu Zhu (IBM), Yunyao Li (IBM), Laura Chiticariu (IBM), Marina Danilevsky (IBM), Sanjana Sahayaraj (IBM)

Wednesday, August 7th, 2019

Deep Learning for NLP with TensorFlow: Cesar Ilharco Magalhaes (Google), Gabriel Ilharco Magalhaes (Google), Jason Baldridge (Google)

Deep Learning at Scale on Databricks: Amir Issaei (Databricks), Brooke Wenig (Databricks)

Deep learning for time series forecasting: Yijing Chen (Microsoft), Dmitry Pechyoni (Microsoft), Angus Taylor (Microsoft), Vanja Paunic (Microsoft)

Introduction to Kubeflow Pipelines: Dan Anghel (Google), Michal Jastrzebski (GitHub), Hamel Husain (GitHub)

Learning Graph Neural Networks with Deep Graph Library: Minjie Wang (New York University), Lingfan Yu (New York University), Da Zheng (AWS), Nick Choma (New York University)

Thursday, August 8th, 2019

From Shallow to Deep Language Representations: Pre-training, Fine-tuning, and Beyond: Aston Zhang (Amazon AI), Sheng Zha (Amazon AI), Haibin Lin (Amazon AI), Alexander Smola (Amazon AI), Leonard Lausen (Amazon AI), Mu Li (Amazon AI), Chenguang Wang (Amazon AI)

From Graph to Knowledge Graph: Mining Large-scale Heterogeneous Networks Using Spark: Zhihong Shen (Microsoft), Yuxiao Dong (Microsoft), Chiyuan Huang (Microsoft), Chieh-Han Wu (Microsoft), Junjie Qian (Microsoft), Anshul Kanakia (Microsoft)

Practice of Efficient Data Collection via Crowdsourcing at Large-Scale: Alexey Drutsa (Yandex), Valentina Fedorova (Yandex), Olga Megorskaya (Yandex), Evfrosiniya Zerminova (Yandex)

Building production-ready recommendation systems at scale: Le Zhang (Microsoft), Tao Wu (Microsoft), Xing Xie (Microsoft), Andreas Argyriou (Microsoft), Ilia Karmanov (Microsoft), Jianxun Lian (Microsoft)

INVITED TALKS

APPLIED DATA SCIENCE

Chairs: Hui Xiong (Baidu/Rutgers University), Sofus Makassy (HackerRank), Johannes Gehrke (Microsoft) and Deepak Argwal (LinkedIn)

Cook Room, Street Level, Egan Center

Tuesday, August 6th, 2019

10:00 AM – 12:00 PM

- **Carlos Guestrin** (Apple). 4 Perspectives in Human-Centered Machine Learning.
- **Ya Xu** (LinkedIn). Evolving Data Science at LinkedIn.
- **Ashok Srivastava** (Intuit). AI for Small Businesses and Consumers: Applications and Innovations.

1:30 PM – 3:30 PM

- **Rama Nemani** (NASA). Earth observations from a new generation of geostationary satellites.
- **Olga Liakhovich** and **Gabriel Domínguez Conde** (Microsoft). Preventing Rhino Poaching through Machine Learning.
- **Ruslan Salakhutdinov** (Apple). Integrating Domain-Knowledge into Deep Learning.
- **Elena Grewal** (Airbnb). Seven Years of Data Science at Airbnb.

4:00 PM – 6:00 PM

- **Jieping Ye** (Didi). Transforming Transportation: A Data Driven Approach.
- **Kumar Chellapilla** (Lyft). Building a better self-driving car: hardware, software, and knowledge.
- **Jairam Ranganathan** (Keeptruckin) Spinning the AI Pinwheel.

Wednesday, August 7th, 2019

10:00 AM – 12:00 PM

- **Neel Sundaresan** (Microsoft). From Code to Data: AI at Scale for Developer Productivity.
- **Ganesh Thondikulam** (Kaiser Permanente). Analytics Journey Map: An approach enable to ML at scale.
- **Paige Maas** (Facebook). Facebook Disaster Maps: Aggregate Insights for Crisis Response & Recovery.

1:30 PM- 3:30 PM

- **David Heckerman** (Amazon). Exploiting high dimensionality in big data.
- **Rich Caruana** (Microsoft). Friends Don't Let Friends Deploy Black-Box Models: The Importance of Intelligibility in Machine Learning.
- **Rómer Rosales** (LinkedIn). Product Ecosystem Optimization at LinkedIn.

Thursday, August 8th, 2019

10:00 AM – 12:00 PM

- **Hongxia Yang** (Alibaba). AliGraph: A Comprehensive Graph Neural Network Platform.
- **Sreenivas Gollapudi** (Google Research). Welfare Maximization in Online Two-sided Marketplaces.
- **Hassan Sawaf** (Amazon Web Services). Applications of AI/ML in Established and New Industries.

1:30 PM – 3:30 PM

- **Joseph Bradley** (Databricks). Addressing Challenges in Data Science: Scale, Skill Sets and Complexity.
- **Konstantinos Katsiapis** (Google). Towards ML Engineering with TensorFlow Extended (TFX).
- **Yinglong Xia** (Huawei). Role of Unified Graph Analysis Platforms for Smart Cities.

APPLIED DATA SCIENCE AND RESEARCH TRACK PAPER ORAL PRESENTATIONS

ADS Program Chairs: Ying Li (Giving Tech Labs) and Romer Rosales (LinkedIn)
Research Program Chairs: George Karypis (University of Minnesota) and Evimaria Terzi (Boston University)

Tuesday, August 6th, 2019 10:00AM - 12:00PM

Applied Data Science Track Session ADS1: Auto-ML and Development Frameworks – Summit 1, Ground Level, Egan Center
Chair: Gabor Melli (Sony PlayStation)

Auto-Keras: An Efficient Neural Architecture Search System
Haifeng Jin (Texas A&M University); Qingquan Song (Texas A&M University); Xia Hu (Texas A&M University)

Pythia: AI-assisted Code Completion System
Alexey Svyatkovskiy (Microsoft); Ying Zhao (Microsoft); Shengyu Fu (Microsoft); Neel Sundaresan (Microsoft)

TF-Ranking: Scalable TensorFlow Library for Learning-to-Rank
Rama Kumar Pasumarthi (Google); Sebastian Bruch (Google); Xuanhui Wang (Google); Cheng Li (Google); Michael Bendersky (Google); Marc Najork (Google); Jan Pfeifer (Google); Nadav Golbandi (Google); Rohan Anil (Google); Stephan Wolf (Google)

Shrinkage Estimators in Online Experiments
Drew Dimmery (Facebook); Eytan Bakshy (Facebook); Jasjeet Sekhon (University of California, Berkeley)

FDML: A Collaborative Machine Learning Framework for Distributed Features
Yaochen Hu (University of Alberta); Di Niu (University of Alberta); Jianming Yang (Tencent); Shengping Zhou (Tencent)

Research Track Session RT1: Neural Networks – Summit 2, Ground Level, Egan Center
Chair: Nesreen Ahmed

Estimating Node Importance in Knowledge Graphs Using Graph Neural Networks
Namyong Park (Carnegie Mellon University & Amazon); Andrey Kan (Amazon); Xin Luna Dong (Amazon); Tong Zhao (Amazon); Christos Faloutsos (Carnegie Mellon University & Amazon)

Certifiable Robustness and Robust Training for Graph Convolutional Networks
Daniel Zügner (Technical University of Munich); Stephan Günnemann (Technical University of Munich)

Effective and Efficient Sports Play Retrieval with Deep Representation Learning
Zheng Wang (Nanyang Technological University); Cheng Long (Nanyang Technological University); Gao Cong (Nanyang Technological University); Ce Ju (Intelligent Driving Group, Baidu Inc.)

Multiple Relational Attention Network for Multi-task Learning
Jiejie Zhao (Beihang University); Bowen Du (Beihang University); Leilei Sun (Beihang University); Fuzhen Zhuang (University of Chinese Academy of Sciences); Weifeng Lv (Beihang University); Hui Xiong (Rutgers University)

The Impact of Person-Organization Fit on Talent Management: A Structure-Aware Convolutional Neural Network Approach
Ying Sun (Institute of Computing Technology, CAS, Baidu Talent Intelligence Center, Baidu Inc., University of Chinese Academy of Sciences); Fuzhen Zhuang (Institute of Computing Technology, CAS & University of Chinese Academy of Sciences); Hengshu Zhu (Baidu Talent Intelligence Center, Baidu Inc.); Xin Song (Baidu Talent Intelligence Center, Baidu Inc.); Qing He (Institute of Computing Technology, CAS & University of Chinese Academy of Sciences); Hui Xiong (Baidu Talent Intelligence Center, Baidu Inc. & Business Intelligence Lab, Baidu Research)

CoSTCo: A Neural Tensor Completion Model for Sparse Tensors
Hanpeng Liu (University of Southern California); Yaguang Li (University of Southern California); Michael Tsang (University of Southern California); Yan Liu (University of Southern California)

Research Track Session RT2: Analyzing Sequential and Temporal Data – Summit 3, Ground Level, Egan Center
Chair: Manuel Gomez Rodriguez

Deep Landscape Forecasting for Real-time Bidding Advertising
Kan Ren (Shanghai Jiao Tong University); Jiarui Qin (Shanghai Jiao Tong University); Lei Zheng (Shanghai Jiao Tong University); Zhengyu Yang (Shanghai Jiao Tong University); Weinan Zhang (Shanghai Jiao Tong University); Yong Yu (Shanghai Jiao Tong University)

Predicting Path Failure In Time-Evolving Graphs
Jia Li (The Chinese University of Hong Kong); Zhichao Han (The Chinese University of Hong Kong); Hong Cheng (The Chinese University of Hong Kong); Jiao Su (The Chinese University of Hong Kong); Pengyun Wang (Noah's Ark Lab, Huawei Technologies); Jianfeng Zhang (Noah's Ark Lab, Huawei Technologies); Lujia Pan (Noah's Ark Lab, Huawei Technologies)

Pairwise Comparisons with Flexible Time-Dynamics
Lucas Maystre (Spotify); Victor Kristof (EPFL); Matthias Grossglauser (EPFL)

Modeling Extreme Events in Time Series Prediction
Daizong Ding (Fudan University); Mi Zhang (Fudan University); Xudong Pan (Fudan University); Min Yang (Fudan University); Xiangnan He (University of Science and Technology of China)

Adversarial Substructured Representation Learning for Mobile User Profiling
Pengyang Wang (Missouri University of Science and Technology); Yanjie Fu (Missouri University of Science and Technology); Hui Xiong (Rutgers University); Xiaolin li (Nanjing University)

Research Track Session RT3: Algorithmic Techniques – Summit 4, Ground Level, Egan Center
Chair: Lingfei Wu

Paper Matching with Local Fairness Constraints
Ari Kobren (University of Massachusetts Amherst); Barna Saha (University of Massachusetts Amherst); Andrew McCallum (University of Massachusetts Amherst)

A Memory-Efficient Sketch Method for Estimating High Similarities in Streaming Sets
Pinghui Wang (Xi'an Jiaotong University); Yiyan Qi (Xi'an Jiaotong University); Yuanming Zhang (Xi'an Jiaotong University); Qiaozhu Zhai (Xi'an Jiaotong University); Chenxu Wang (Xi'an Jiaotong University); John C.S. Lui (The Chinese University of Hong Kong); Xiaohong Guan (Xi'an Jiaotong University & Tsinghua University)

Revisiting kd-tree for Nearest Neighbor Search
Parikshit Ram (IBM Research AI); Kaushik Sinha (Wichita State University)

Adversarially Robust Submodular Maximization under Knapsack Constraints
Dmitrii Avdiukhin (Indiana University); Slobodan Mitrovic (Massachusetts Institute of Technology); Grigory Yaroslavtsev (Indiana University); Samson Zhou (Indiana University)

MinJoin: Efficient Edit Similarity Joins via Local Hash Minima
Haoyu Zhang (Indiana University Bloomington); Qin Zhang (Indiana University Bloomington)

Coresets for Minimum Enclosing Balls over Sliding Windows
Yanhao Wang (National University of Singapore); Yuchen Li (Singapore Management University); Kian-Lee Tan (National University of Singapore)

1:30PM - 3:30PM

Applied Data Science Track Session ADS2: Language Models and Text Mining – Summit 1, Ground Level, Egan Center
Chair: Elena Baralis (Politecnico di Torino)

Automatic Dialogue Summary Generation for Customer Service
Chunyi Liu (AI Labs, Didi Chuxing); Peng Wang (AI Labs, Didi Chuxing); Jiang Xu (AI Labs, Didi Chuxing); Zang Li (AI Labs, Didi Chuxing); Jieping Ye (AI Labs, Didi Chuxing)

Detection of Review Abuse via Semi-Supervised Binary Multi-Target Tensor Decomposition
Anil R Yelundur (Amazon); Vineet Chaoji (Amazon); Bamdev Mishra (Microsoft India)

Unsupervised Clinical Language Translation
Wei-Hung Weng (Massachusetts Institute of Technology); Yu-An Chung (Massachusetts Institute of Technology); Peter Szolovits (Massachusetts Institute of Technology)

Gmail Smart Compose: Real-Time Assisted Writing
Mia Xu Chen (Google); Benjamin N. Lee (Google); Gagan Bansal (Google); Yuan Cao (Google); Shuyuan Zhang (Google); Justin Lu (Google); Jackie Tsay (Google); Yinan Wang (Google); Andrew M. Dai (Google); Zhifeng Chen (Google); Timothy Sohn (Google); Yonghui Wu (Google)

Naranjo Question Answering using End-to-End Multi-task Learning Model
Bhanu Pratap Singh Rawat (University of Massachusetts Amherst); Fei Li (University of Massachusetts Lowell); Hong Yu (University of Massachusetts Lowell)

Applied Data Science Track Session ADS3: Urbanism and Mobility – Summit 5/6, Ground Level, Egan Center
Chair: Mohak Shah (LG Electronics)

DeepUrbanEvent: A System for Predicting Citywide Crowd Dynamics at Big Events
Renhe Jiang (The University of Tokyo & National Institute of Advanced Industrial Science and Technology); Xuan Song (The University of Tokyo & National Institute of Advanced Industrial Science and Technology); Dou Huang (The University of Tokyo); Xiaoya Song (The University of Tokyo & Harbin Institute of Technology); Tianqi Xia (The University of Tokyo & National Institute of Advanced Industrial Science and Technology); Zekun Cai (The University of Tokyo); Zhaonan Wang (National Institute of Advanced Industrial Science and Technology); Kyoung-Sook Kim (National Institute of Advanced Industrial Science and Technology); Ryosuke Shibasaki (The University of Tokyo)

Hydra: A Personalized and Context-Aware Multi-Modal Transportation Recommendation System
Hao Liu (Baidu Research); Yongxin Tong (Beihang University); Panpan Zhang (Baidu Research); Xinjiang Lu (Baidu Research); Jianguo Duan (Baidu Research); Hui Xiong (Business Intelligence Lab)

A Deep Value-network Based Approach for Multi-Driver Order Dispatching
Xiaocheng Tang (AI Labs, Didi Chuxing); Zhiwei (Tony) Qin (AI Labs, Didi Chuxing); Fan Zhang (AI Labs, Didi Chuxing); Zhaodong Wang (Washington State University); Zhe Xu (Didi Chuxing); Yintai Ma (Northwestern University); Hongtu Zhu (AI Labs, Didi Chuxing); Jieping Ye (AI Labs, Didi Chuxing)

Hard to Park? Estimating Parking Difficulty at Scale
Neha Arora (Google Research); James Cook (Google Research); Ravi Kumar (Google Research); Ivan Kuznetsov (Google Research); Yechen Li (Google Research); Huai-Jen Liang (Google Research); Andrew Miller (Google Research); Andrew Tomkins (Google Research); Iveel Tsogsuren (Google Research); Yi Wang (Google Research)

Nostalgin: Extracting 3D City Models from Historical Image Data
Amol Kapoor (Google Research); Hunter Larco (Google Research); Raimondas Kiveris (Google Research)

Research Track Session RT4: Embeddings I – Summit 2, Ground Level, Egan Center
Chair: Kanishka Bhaduri

ProGAN: Network Embedding via Proximity Generative Adversarial Network
Hongchang Gao (University of Pittsburgh); Jian Pei (Simon Fraser University); Heng Huang (University of Pittsburgh)

Scalable Global Alignment Graph Kernel Using Random Features: From Node Embedding to Graph Embedding
Lingfei Wu (IBM Research); Ian En-Hsu Yen (Carnegie Mellon University); Zhen Zhang (Washington University in St. Louis); Kun Xu (IBM Research); Liang Zhao (George Mason University); Xi Peng (University of Delaware); Yinglong Xia (Huawei); Charu Aggarwal (IBM Research)

Scalable Graph Embeddings via Sparse Transpose Proximities
Yuan Yin (Renmin University of China); Zhewei Wei (Renmin University of China)

Enhancing Domain Word Embedding via Latent Semantic Imputation
Shibo Yao (New Jersey Institute of Technology); Dantong Yu (New Jersey Institute of Technology); Keli Xiao (Stony Brook University)

EpiDeep: Exploiting Embeddings for Epidemic Forecasting
Bijaya Adhikari (Virginia Tech); Xinfeng Xu (Virginia Tech); Naren Ramakrishnan (Virginia Tech); B. Aditya Prakash (Virginia Tech)

Research Track Session RT5: Privacy and Policy Learning – Summit 3, Ground Level, Egan Center
Chair: Jie Tang

PrivPy: General and Scalable Privacy-Preserving Data Mining
Yi Li (Tsinghua University); Wei Xu (Tsinghua University)

Auditing Data Provenance in Text-Generation Models
Congzheng Song (Cornell University); Vitaly Shmatikov (Cornell Tech)

Focused Context Balancing for Robust Offline Policy Evaluation
Hao Zou (Tsinghua University & Beijing National Research Center for Information Science and Technology (BNRist).); Kun Kuang (Tsinghua University); Boqi Chen (Boston University & Tsinghua); Peixuan Chen (Tencent); Peng Cui (Tsinghua University)

Off-policy Learning for Multiple Loggers
Li He (JD.com); Long Xia (JD.com); Wei Zeng (Institute of Computing Technology, CAS); Zhi-Ming Ma (Academy of Mathematics and Systems Science, CAS); Yihong Zhao (JD.com); Dawei Yin (JD.com)

Figuring out the User in a Few Steps: Bayesian Multifidelity Active Search with Cokriging
Nikita Klyuchnikov (Skoltech); Davide Mottin (Aarhus University); Georgia Koutrika (Athena Research and Innovation Center); Emmanuel Möller (University of Bonn); Panagiotis Karras (Aarhus University)

Research Track Session RT6: Network Science – Summit 4, Ground Level, Egan Center
Chair: Danai Koutra

Network Density of States
Kun Dong (Cornell University); Austin R. Benson (Cornell University); David Bindel (Cornell University)

Link Prediction with Signed Latent Factors in Signed Social Networks
Pinghua Xu (Wuhan University & Macquarie University); Wenbin Hu (Wuhan University & Shenzhen Research Institute, Wuhan University); Jia Wu (Macquarie University); Bo Du (Wuhan University)

Attribute-Driven Backbone Discovery
Sheng Guan (Washington State University); Hanchao Ma (Washington State University); Yinghui Wu (Washington State University & Pacific Northwest National Laboratory)

Fates of Microscopic Social Ecosystems: Keep Alive or Dead?
Haoyang Li (Tsinghua University); Peng Cui (Tsinghua University); Chengxi Zang (Tsinghua University); Tianyang Zhang (Tsinghua University); Wenwu Zhu (Tsinghua University); Yishi Lin (Tencent)

MCNE: An End-to-End Framework for Learning Multiple Conditional Network Representations of Social Network
Hao Wang (University of Science and Technology of China); Tong Xu (University of Science and Technology of China); Qi Liu (University of Science and Technology of China); Defu Lian (University of Science and Technology of China); Enhong Chen (University of Science and Technology of China); Dongfang Du (Tencent Inc); Han Wu (University of Science and Technology of China); Wen Su (Tencent Inc)

Learning Dynamic Context Graphs for Predicting Social Events
Songgaojun Deng (Stevens Institute of Technology); Huzefa Rangwala (George Mason University); Yue Ning (Stevens Institute of Technology)

4:00PM - 6:00PM

Applied Data Science Track Session ADS4: Real-Time and Online – Summit 1, Ground Level, Egan Center
Chair: Ying Shan (Tencent)

Time-Series Anomaly Detection Service at Microsoft
Hansheng Ren (Microsoft); Bixiong Xu (Microsoft); Yujing Wang (Microsoft); Chao Yi (Microsoft); Congrui Huang (Microsoft); Xiaoyu Kou (Microsoft); Tony Xing (Microsoft); Mao Yang (Microsoft); Jie Tong (Microsoft); Qi Zhang (Microsoft)

Real-time On-Device Troubleshooting Recommendation for Smartphones
Keiichi Ochiai (NTT DOCOMO, INC.); Kohei Senkawa (NTT DOCOMO, INC.); Naoki Yamamoto (NTT DOCOMO, INC.); Yuya Tanaka (NTT DOCOMO, INC.); Yusuke Fukazawa (NTT DOCOMO, INC.)

Real-time Attention Based Look-alike Model for Recommender System
Yudan Liu (WeiXin Group, Tencent Inc.); Kaikai Ge (WeiXin Group, Tencent Inc.); Xu Zhang (WeiXin Group, Tencent Inc.); Leyu Lin (WeiXin Group, Tencent Inc.)

Anomaly Detection for an E-commerce Pricing System
Jagdish Ramakrishnan (Walmart Labs); Elham Shaabani (Walmart Labs); Chao Li (Walmart Labs); Matyas A. Sustik (Walmart Labs)

Online Amnestic DTW to allow Real-Time Golden Batch Monitoring
Chin-Chia Michael Yeh (University of California, Riverside); Yan Zhu (University of California, Riverside); Hoang Anh Dau (University of California, Riverside); Amirali Darvishzadeh (University of California, Riverside); Mikhail Noskov (Aspen Technology); Eamonn Keogh (University of California, Riverside)

Research Track Session RT7: Graph Neural Networks – Summit 2, Ground Level, Egan Center
Chair: Jiliang Tang

Conditional Random Field Enhanced Graph Convolutional Neural Networks
Hongchang Gao (University of Pittsburgh & JD Finance America Corporation); Jian Pei (Simon Fraser University); Heng Huang (University of Pittsburgh & JD Finance America Corporation)

Robust Graph Convolutional Networks Against Adversarial Attacks
Dingyuan Zhu (Tsinghua University); Ziwei Zhang (Tsinghua University); Peng Cui (Tsinghua University); Wenwu Zhu (Tsinghua University)

GCN-MF: Disease-Gene Association Identification By Graph Convolutional Networks and Matrix Factorization
Peng Han (King Abdullah University of Science and Technology); Peng Yang (Cognitive Computing Lab, Baidu Research USA); Peilin Zhao (Tencent AI Lab); Shuo Shang (University of Electronic Science and Technology of China & Inception Institute of Artificial Intelligence); Yong Liu (Alibaba-NTU Singapore Joint Research Institute, Nanyang Technological University); Jiayu Zhou (Michigan State University); Xin Gao (King Abdullah University of Science and Technology); Panos Kalnis (King Abdullah University of Science and Technology)

Cluster-GCN: An Efficient Algorithm for Training Deep and Large Graph Convolutional Networks
Wei-Lin Chiang (National Taiwan University & Google Research); Xuanqing Liu (University of California, Los Angeles & Google Research); Si Si (Google Research); Yang Li (Google Research); Samy Bengio (Google Research); Cho-Jui Hsieh (University of California, Los Angeles)

Graph Representation Learning via Hard and Channel-Wise Attention Networks
Hongyang Gao (Texas A&M University); Shuiwang Ji (Texas A&M University)

Origin-Destination Matrix Prediction via Graph Convolution: A New Perspective of Passenger Demand Modeling
Yuandong Wang (Beihang University); Hongzhi Yin (The University of Queensland); Hongxu Chen (The University of Queensland); Tianyu Wo (Beihang University); Jie Xu (University of Leeds); Kai Zheng (University of Electronic Science and Technology)

Research Track Session RT8: Knowledge Extraction – Summit 3, Ground Level, Egan Center
Chair: Huan Sun

Universal Representation Learning of Knowledge Bases by Jointly Embedding Instances and Ontological Concepts
Junheng Hao (University of California Los Angeles); Muhao Chen (University of California Los Angeles); Wenchao Yu (University of California Los Angeles); Yizhou Sun (University of California Los Angeles); Wei Wang (University of California Los Angeles)

Relation Extraction via Domain-aware Transfer Learning
Shimin Di (The Hong Kong University of Science and Technology); Yanyan Shen (Shanghai Jiao Tong University); Lei Chen (The Hong Kong University of Science and Technology)

Exploiting Cognitive Structure for Adaptive Learning
Qi Liu (University of Science and Technology of China); Shiwei Tong (University of Science and Technology of China); Chuanren Liu (University of Tennessee); Hongke Zhao (Tianjin University); Enhong Chen (University of Science and Technology of China); Haiping Ma (iFLYTEK CO., LTD & State Key Laboratory of Cognitive Intelligence); Shijin Wang (iFLYTEK CO., LTD & State Key Laboratory of Cognitive Intelligence)

Mining Algorithm Roadmap in Scientific Publications
Hanwen Zha (University of California, Santa Barbara); Wenhu Chen (University of California, Santa Barbara); Keqian Li (University of California, Santa Barbara); Xifeng Yan (University of California, Santa Barbara)

Knowledge-aware Graph Neural Networks with Label Smoothness Regularization for Recommender Systems
Hongwei Wang (Stanford University); Fuzheng Zhang (Meituan-Dianping Group); Mengdi Zhang (Meituan-Dianping Group); Jure Leskovec (Stanford University); Miao Zhao (Hong Kong Polytechnic University); Wenjie Li (Hong Kong Polytechnic University); Zhongyuan Wang (Meituan-Dianping Group)

Adaptive Graph Guided Disambiguation for Partial Label Learning
Deng-Bao Wang (Southwest University); Li Li (Southwest University); Min-Ling Zhang (Southeast University & Ministry of Education)

Research Track Session RT9: Mining in Emerging Applications I – Summit 4, Ground Level, Egan Center
Chair: Shandian Zhe

SurfCon: Synonym Discovery on Privacy-Aware Clinical Data
Zhen Wang (The Ohio State University); Xiang Yue (The Ohio State University); Soheil Moosavinasab (Abigail Wexner Research Institute at Nationwide Childrenés Hospital); Yungui Huang (Abigail Wexner Research Institute at Nationwide Childrenés Hospital); Simon Lin (Abigail Wexner Research Institute at Nationwide Childrenés Hospital); Huan Sun (The Ohio State University)

Hierarchical Multi-Task Word Embedding Learning for Synonym Prediction
Hongliang Fei (Baidu Research); Shulong Tan (Baidu Research); Ping Li (Baidu Research)

GroupINN: Grouping-based Interpretable Neural Network for Classification of Limited, Noisy Brain Data
Yujun Yan (University of Michigan); Jiong Zhu (University of Michigan); Marlena Duda (University of Michigan); Eric Solarz (University of Michigan); Chandra Sripada (University of Michigan); Danai Koutra (University of Michigan)

PerDREP: Personalized Drug Effectiveness Prediction from Longitudinal Observational Data
Sanjoy Dey (IBM T. J. Watson Research Center); Ping Zhang (Ohio State University); Daby Sow (IBM T. J. Watson Research Center); Kenney Ng (IBM T. J. Watson Research Center)

Retaining Privileged Information for Multi-Task Learning
Fengyi Tang (Michigan State University); Cao Xiao (IQVIA); Fei Wang (Weill Cornell Medical College); Jiayu Zhou (Michigan State University); Li-wei H. Lehman (Massachusetts Institute of Technology)

A Permutation Approach to Assess Confounding in Machine Learning Applications for Digital Health
Elias Chaibub Neto (Sage Bionetworks); Abhishek Pratap (Sage Bionetworks); Thanneer M. Perumal (Sage Bionetworks); Meghasyam Tummalacherla (Sage Bionetworks); Brian M. Bot (Sage Bionetworks); Lara Mangravite (Sage Bionetworks); Larsson Omberg (Sage Bionetworks)

Wednesday August 7th, 2019 10:00AM - 12:00PM

Applied Data Science Track Session ADS5: Scalability and Novel Applications – Summit 1, Ground Level, Egan Center
Chair: Marc Najork (Google)

Large-Scale Training Framework for Video Annotation
Seong Jae Hwang (University of Wisconsin-Madison); Joonseok Lee (Google Research); Balakrishnan Varadarajan (Google Research); Ariel Gordon (Google Research); Zheng Xu (Google Research); Apostol (Paul) Natsev (Google Research)

A Data-Driven Approach for Multi-level Packing Problems in Manufacturing Industry
Lei Chen (Huawei Noah's Ark Lab); Xialiang Tong (Huawei Noah's Ark Lab); Mingxuan Yuan (Huawei Noah's Ark Lab); Jia Zeng (Huawei Noah's Ark Lab); Lei Chen (Hong Kong University of Science and Technology)

Actions Speak Louder than Goals: Valuing Player Actions in Soccer
Tom Decroos (KU Leuven); Lotte Bransen (SciSports); Jan Van Haaren (SciSports); Jesse Davis (KU Leuven)

Enabling Onboard Detection of Events of Scientific Interest for the Europa Clipper Spacecraft
Kiri L. Wagstaff (California Institute of Technology); Gary Doran (California Institute of Technology); Ashley Davies (California Institute of Technology); Saadat Anwar (Arizona State University); Srija Chakraborty (Arizona State University); Marissa Cameron (California Institute of Technology); Ingrid Daubar (California Institute of Technology); Cynthia Phillips (California Institute of Technology)

DeepRoof: A Data-driven Approach For Solar Potential Estimation Using Rooftop Imagery
Stephen Lee (University of Massachusetts, Amherst); Srinivasan Iyengar (Microsoft Research, Bangalore); Menghong Feng (University of Massachusetts, Amherst); Prashant Shenoy (University of Massachusetts, Amherst); Subhransu Maji (University of Massachusetts, Amherst)

Research Track Session RT10: Embeddings II – Summit 2, Ground Level, Egan Center
Chair: Jundong Li

Individualized Indicator for All: Stock-wise Technical Indicator Optimization with Stock Embedding
Zhige Li (Shanghai Jiao Tong University); Derek Yang (Tsinghua University); Li Zhao (Microsoft); Jiang Bian (Microsoft); Tao Qin (Microsoft); Tie-Yan Liu (Microsoft)

Efficient Global String Kernel with Random Features: Beyond Counting Substructures
Lingfei Wu (IBM Research); Ian En-Hsu Yen (CMU); Siyu Huo (IBM Research); Liang Zhao (George Mason University); Kun Xu (IBM Research); Liang Ma (IBM Research); Shouling Ji (Zhejiang University); Charu Aggarwal (IBM Research)

HATS: A Hierarchical Sequence-Attention Framework for Inductive Set-of-Sets Embeddings
Changping Meng (Purdue University); Jiasen Yang (Purdue University); Bruno Ribeiro (Purdue University); Jennifer Neville (Purdue University)

TUBE: Embedding Behavior Outcomes for Predicting Success
Daheng Wang (University of Notre Dame); Tianwen Jiang (University of Notre Dame & Harbin Institute of Technology); Nitesh V. Chawla (University of Notre Dame); Meng Jiang (University of Notre Dame)

Multi-Relational Classification via Bayesian Ranked Non-Linear Embeddings
Ahmed Rashed (University of Hildesheim); Josif Grabocka (University of Hildesheim); Lars Schmidt-Thieme (University of Hildesheim)

Learning Network-to-Network Model for Content-rich Network Embedding
Zhicheng He (Nankai University); Jie Liu (Nankai University); Na Li (Nankai University); Yalou Huang (Nankai University)

Research Track Session RT11: Clustering and Visualization – Summit 3, Ground Level, Egan Center
Chair: David Anastasiu

Scalable Hierarchical Clustering with Tree Grafting
Nicholas Monath (University of Massachusetts Amherst); Ari Kobren (University of Massachusetts Amherst); Akshay Krishnamurthy (Microsoft Research); Michael Glass (International Business Machines); Andrew McCallum (University of Massachusetts Amherst)

K-Multiple-Means: A Multiple-Means Clustering Method with Specified K Clusters
Feiping Nie (School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University); Cheng-Long Wang (School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University); Xuelong Li (School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University)

A Multiscale Scan Statistic for Adaptive Submatrix Localization
Yuchao Liu (Microsoft Corporation); Ery Arias-Castro (University of California San Diego)

Robust Task Grouping with Representative Tasks for Clustered Multi-Task Learning
Yaqiang Yao (University of Science and Technology of China); Jie Cao (Nanjing University of Finance and Economics); Huanhuan Chen (University of Science and Technology of China)

AtSNE: Efficient and Robust Visualization on GPU through Hierarchical Optimization
Cong Fu (Zhejiang University); Yonghui Zhang (Zhejiang University); Deng Cai (Alibaba-Zhejiang University Joint Institute of Frontier Technologies); Xiang Ren (University of Southern California)

Research Track Session RT12: Recommender Systems I – Summit 4, Ground Level, Egan Center
Chair: Xia Ning

Predicting Embedding Trajectories in Temporal Interaction Networks
Srijan Kumar (Stanford University); Xikun Zhang (University of Illinois); Jure Leskovec (Stanford University)

Environment Reconstruction with Hidden Confounders for Reinforcement Learning based Recommendation
Wenjie Shang (Nanjing University); Yang Yu (Nanjing University); Qingyang Li (AI Labs, Didi Chuxing); Zhiwei Qin (AI Labs, Didi Chuxing); Yiping Meng (AI Labs, Didi Chuxing); Jieping Ye (AI Labs, Didi Chuxing)

Empowering A* Search Algorithms with Neural Networks for Personalized Route Recommendation
Jingyuan Wang (Beihang University); Ning Wu (Beihang University); Wayne Xin Zhao (Renmin University of China); Fanzhang Peng (Beihang University); Xin Lin (Beihang University)

Effective and Efficient Reuse of Past Travel Behavior for Route Recommendation
Lisi Chen (Inception Institute of Artificial Intelligence); Shuo Shang (University of Electronic Science and Technology of China & Inception Institute of Artificial Intelligence); Christian S. Jensen (Aalborg University); Bin Yao (Shanghai Jiao Tong University); Zhiwei Zhang (Hong Kong Baptist University); Ling Shao (Inception Institute of Artificial Intelligence)

State-Sharing Sparse Hidden Markov Models for Personalized Sequences
Hongzhi Shi (Tsinghua University); Chao Zhang (Georgia Institute of Technology); Quanming Yao (4Paradigm Inc.); Yong Li (Tsinghua University); Funing Sun (Tencent Inc.); Depeng Jin (Tsinghua University)

1:30PM - 3:30PM

Applied Data Science Track Session ADS6: Environment and Sustainability – Summit 1, Ground Level, Egan Center
Chair: Noam Koenigstein (Microsoft)

Nonparametric Mixture of Sparse Regressions on Spatio-Temporal Data -- An Application to Climate Prediction
Yumin Liu (Northeastern University); Junxiang Chen (Northeastern University); Auroop Ganguly (Northeastern University); Jennifer Dy (Northeastern University)

Deep Uncertainty Quantification: A Machine Learning Approach for Weather Forecasting
Bin Wang (University of Technology Sydney & Southwest Jiaotong University); Jie Lu (University of Technology Sydney); Zheng Yan (University of Technology Sydney); Huaishao Luo (Southwest Jiaotong University); Tianrui Li (Southwest Jiaotong University); Yu Zheng (Xidian University & JD Intelligent Cities Research); Guangquan Zhang (University of Technology Sydney)

Towards Sustainable Dairy Management - A Machine Learning Enhanced Method for Estrus Detection
Kevin Fauvel (Universite Rennes, Inria, CNRS, IRISA); Véronique Masson (Universite Rennes, Inria, CNRS, IRISA); Élisa Fromont (Universite Rennes, Inria, CNRS, IRISA); Philippe Faverdin (PEGASE, INRA, AGROCAMPUS OUEST); Alexandre Termier (Universite Rennes, Inria, CNRS, IRISA)

Precipitation Nowcasting with Satellite Imagery
Vadim Lebedev (Yandex); Vladimir Ivashkin (Yandex); Irina Rudenko (Yandex); Alexander Ganshin (Yandex); Alexander Molchanov (Yandex); Sergey Ovcharenko (Yandex); Ruslan Grokhovetskiy (Yandex); Ivan Bushmarinov (Yandex); Dmitry Solomentsev (Yandex)

AccuAir: Winning Solution to Air Quality Prediction for KDD Cup 2018
Zhipeng Luo (DeepBlue Technology); Jianqiang Huang (Peking University); Ke Hu (Alibaba Group); Xue Li (Microsoft); Peng Zhang (Tianjin University)

Research Track Session RT13: Learning – Summit 2, Ground Level, Egan Center
Chair: Xiang Ren

Learning from Incomplete and Inaccurate Supervision
Zhen-Yu Zhang (Nanjing University); Peng Zhao (Nanjing University); Yuan Jiang (Nanjing University); Zhi-Hua Zhou (Nanjing University)

Automating Feature Subspace Exploration via Multi-Agent Reinforcement Learning
Kunpeng Liu (University of Central Florida); Yanjie Fu (University of Central Florida); Pengfei Wang (CNIC, Chinese Academy of Sciences); Le Wu (Hefei University of Technology); Rui Bo (Missouri Univ. of Sci. and Tech.); Xiaolin Li (Nanjing University)

Task-Adversarial Co-Generative Nets
Pei Yang (Arizona State University & South China University of Technology); Qi Tan (South China Normal University); Hanghang Tong (Arizona State University); Jingrui He (Arizona State University)

Learning Class-Conditional GANs with Active Sampling
Ming-Kun Xie (Nanjing University of Aeronautics and Astronautics); Sheng-Jun Huang (Nanjing University of Aeronautics and Astronautics)

Adaptive Unsupervised Feature Selection on Attributed Networks
Jundong Li (Arizona State University); Ruocheng Guo (Arizona State University); Chenghao Liu (Singapore Management University); Huan Liu (Arizona State University)

Identifiability of Cause and Effect using Regularized Regression
Alexander Marx (Max Planck Institute for Informatics); Jilles Vreeken (CISPA Helmholtz Center for Information Security)

Research Track Session RT14: Anomaly Detection – Summit 3, Ground Level, Egan Center
Chair: Leman Akoglu

EdMot: An Edge Enhancement Approach for Motif-aware Community Detection
Pei-Zhen Li (Sun Yat-sen University); Ling Huang (Sun Yat-sen University); Chang-Dong Wang (Sun Yat-sen University); Jian-Huang Lai (Sun Yat-sen University)

Sequential Anomaly Detection using Inverse Reinforcement Learning
Min-hwan Oh (Columbia University); Garud Iyengar (Columbia University)

Deep Anomaly Detection with Deviation Networks
Guansong Pang (University of Adelaide); Chunhua Shen (University of Adelaide); Anton van den Hengel (University of Adelaide)

Discovering Unexpected Local Nonlinear Interactions in Scientific Black-box Models
Michael Doron (Hebrew University of Jerusalem); Idan Segev (Hebrew University of Jerusalem); Dafna Shahaf (Hebrew University of Jerusalem)

Research Track Session RT15: Mining in Emerging Applications II – Summit 4, Ground Level, Egan Center
Chair: Petko Bogdanov

Optimizing Impression Counts for Outdoor Advertising
Yipeng Zhang (RMIT University); Yuchen Li (Singapore Management University); Zhifeng Bao (RMIT University); Songsong Mo (Wuhan Univeristy & RMIT University); Ping Zhang (Huawei)

Three-Dimensional Stable Matching Problem for Spatial Crowdsourcing Platforms
Boyang Li (Northeastern University); Yurong Cheng (Beijing Institute of Technology); Ye Yuan (Northeastern University); Guoren Wang (Beijing Institute of Technology); Lei Chen (The Hong Kong University of Science and Technology)

Hidden POI Ranking with Spatial Crowdsourcing
Yue Cui (University of Electronic Science and Technology of China); Liwei Deng (University of Electronic Science and Technology of China); Yan Zhao (Soochow University); Bin Yao (Shanghai Jiao Tong University); Vincent W. Zheng (WeBank); Kai Zheng (University of Electronic Science and Technology of China)

Hidden Markov Contour Tree: A Spatial Structured Model for Hydrological Applications
Zhe Jiang (University of Alabama); Arpan Man Sainju (University of Alabama)

Urban Traffic Prediction from Spatio-Temporal Data Using Deep Meta Learning
Zheyi Pan (Shanghai Jiaotong University); Yuxuan Liang (Xidian University); Weifeng Wang (Shanghai Jiaotong University); Yong Yu (Shanghai Jiaotong University); Yu Zheng (JD Intelligent Cities Research); Junbo Zhang (JD Intelligent Cities Research)

Co-Prediction of Multiple Transportation Demands Based on Deep Spatio-Temporal Neural Network
Junchen Ye (Beihang University); Leilei Sun (Beihang University); Bowen Du (Beihang University); Yanjie Fu (University of Central Florida); Xinran Tong (Beihang University); Hui Xiong (Rutgers University)

Thursday August 8th, 2019 10:00AM - 12:00PM

Applied Data Science Track Session ADS7: Entity Extraction, Linking, and Search – Summit 5/6, Ground Level, Egan Center
Chair: Myra Spiliopoulou (Otto-von-Guericke-University Magdeburg)

A Collaborative Learning Framework to Tag Refinement for Points of Interest
Jingbo Zhou (Baidu Research & National Engineering Laboratory of Deep Learning Technology and Application); Shan Gou (Baidu Research & University of Electronic Science and Technology of China); Renjun Hu (Business Intelligence Lab, Baidu Research); Dongxiang Zhang (Zhejiang University); Jin Xu (Baidu Research); Xuehui Wu (University of Electronic Science and Technology of China); Airong Jiang (Baidu Research & National Engineering Laboratory of Deep Learning Technology and Application); Hui Xiong (Baidu Research, National Engineering Laboratory of Deep Learning Technology and Application, & Rutgers University)

Combining Decision Trees and Neural Networks for Learning-to-Rank in Personal Search
Pan Li (University of Illinois at Urbana - Champaign & Google Inc.); Zhen Qin (Google Inc.); Xuanhui Wang (Google Inc.); Donald Metzler (Google Inc.)

Fairness-Aware Ranking in Search & Recommendation Systems with Application to LinkedIn Talent Search
Sahin Cem Geyik (LinkedIn Corporation); Stuart Ambler (LinkedIn Corporation); Krishnaram Kenthapadi (LinkedIn Corporation)

How to Invest my Time: Lessons from Human-in-the-Loop Entity Extraction
Shanshan Zhang (Temple University); Lihong He (Temple University); Eduard Dragut (Temple University); Slobodan Vucetic (Temple University)

OAG: Toward Linking Large-scale Heterogeneous Entity Graphs
Fanjin Zhang (Tsinghua University); Xiao Liu (Tsinghua University); Jie Tang (Tsinghua University); Yuxiao Dong (Microsoft Research); Peiran Yao (Tsinghua University); Jie Zhang (Tsinghua University); Xiaotao Gu (Tsinghua University); Yan Wang (Tsinghua University); Bin Shao (Microsoft Research); Rui Li (Microsoft Research); Kuansan Wang (Microsoft Research)

Applied Data Science Track Session ADS8: Sensor and Consumer Services – Summit 1, Ground Level, Egan Center
Chair: Romer Rosales (LinkedIn)

Towards Identifying Impacted Users in Cellular Services
Shobha Venkataraman (Unaffiliated); Jia Wang (AT&T Labs -- Research)

Ambulatory Atrial Fibrillation Monitoring Using Wearable Photoplethysmography with Deep Learning
Yichen Shen (Samsung Strategy and Innovation Center); Maxime Voisin (Stanford University); Alireza Aliamiri (Samsung Strategy and Innovation Center); Anand Avati (Stanford University); Awni Hannun (Stanford University); Andrew Ng (Stanford University)

Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams
Richard Chen (Apple Inc.); Filip Jankovic (Evidation Health, Inc.); Nikki Marinsek (Evidation Health, Inc.); Luca Foschini (Evidation Health, Inc.); Lampros Kourtis (Evidation Health, Inc.); Alessio Signorini (Evidation Health, Inc.); Melissa Pugh (Eli Lilly and Company); Jie Shen (Eli Lilly and Company); Roy Yaari (Eli Lilly and Company); Vera Maljkovic (Eli Lilly and Company); Marc Sunga (Eli Lilly and Company); Han Hee Song (Apple Inc.); Hyun Joon Jung (Apple Inc.); Belle Tseng (Apple Inc.); Andrew Trister (Apple Inc.)

Sequence Multi-task Learning to Forecast Mental Wellbeing from Sparse Self-reported Data
Dimitris Spathis (University of Cambridge); Sandra Servia-Rodriguez (University of Cambridge); Katayoun Farrahi (University of Southampton); Cecilia Mascolo (University of Cambridge); Jason Rentfrow (University of Cambridge)

Learning to Prescribe Interventions for Tuberculosis Patients Using Digital Adherence Data
Jackson A. Killian (University of Southern California); Bryan Wilder (University of Southern California); Amit Sharma (Microsoft Research India); Vinod Choudhary (RNTCP, Mumbai); Bistra Dilkina (University of Southern California); Milind Tambe (University of Southern California)

Research Track Session RT16: Machine Learning Themes I – Summit 2, Ground Level, Egan Center
Chair: Shandian Zhe

Dynamical Origins of Distribution Functions
Chengxi Zang (Tsinghua University & Weill Cornell Medicine); Peng Cui (Tsinghua University); Wenwu Zhu (Tsinghua University); Fei Wang (Weill Cornell Medicine)

Scaling Multinomial Logistic Regression Via Hybrid Parallelism
Parameswaran Raman (University of California Santa Cruz); Sriram Srinivasan (University of California Santa Cruz); Shin Matsushima (University of Tokyo); Xinhua Zhang (University of Illinios Chicago); Hyokun Yun (Amazon); S.V.N. Vishwanathan (Amazon)

DeepGBM: A Deep Learning Framework Distilled by GBDT for Online Prediction Tasks
Guolin Ke (Microsoft Research); Zhenhui Xu (Peking University); Jia Zhang (Microsoft Research); Jiang Bian (Microsoft Research); Tie-Yan Liu (Microsoft Research)

Graph-based Semi-Supervised & Active Learning for Edge Flows
Junteng Jia (Cornell University); Michael T. Schaub (Massachusetts Institute of Technology & University of Oxford); Santiago Segarra (Rice University); Austin R. Benson (Cornell University)

A Minimax Game for Instance based Selective Transfer Learning
Bo Wang (Alibaba Group); Minghui Qiu (Alibaba Group & Zhejiang University); Xisen Wang (Alibaba Group); Yaliang Li (Alibaba); Yu Gong (Alibaba); Xiaoyi Zeng (Aliabab); Jun Huang (Alibaba); Bo Zheng (Alibaba); Deng Cai (Alibaba); Jingren Zhou (Alibaba)

AutoNE: Hyperparameter Optimization for Massive Network Embedding
Ke Tu (Tsinghua University); Jianxin Ma (Tsinghua University); Peng Cui (Tsinghua University); Jian Pei (Simon Fraser University and JD.com); Wenwu Zhu (Tsinghua University)

Research Track Session RT17: Interpretability – Summit 3, Ground Level, Egan Center
Chair: Liang Zhao

Learning Interpretable Metric between Graphs: Convex Formulation and Computation with Graph Mining
Tomoki Yoshida (Nagoya Institute of Technology); Ichiro Takeuchi (Nagoya Institute of Technology & National Institute for Material Science & RIKEN Center for Advanced Intelligence Project); Masayuki Karasuyama (Nagoya Institute of Technology & National Institute for Material Science & Japan Science and Technology Agency)

Axiomatic Interpretability for Multiclass Additive Models
Xuezhou Zhang (University of Wisconsin-Madison); Sarah Tan (Cornell University); Paul Koch (Microsoft Research); Yin Lou (Ant Financial); Urszula Chajewska (Microsoft); Rich Caruana (Microsoft Research)

Incorporating Interpretability into Latent Factor Models via Fast Influence Analysis
Weiyu Cheng (Shanghai Jiao Tong University); Yanyan Shen (Shanghai Jiao Tong University); Linpeng Huang (Shanghai Jiao Tong University); Yanmin Zhu (Shanghai Jiao Tong University)

Improving the Quality of Explanations with Local Embedding Perturbations
Yunzhe Jia (University of Melbourne); James Bailey (University of Melbourne); Kotagiri Ramamohanarao (University of Melbourne); Christopher Leckie (University of Melbourne); Michael E. Houle (National Institute of Informatics)

Log2Intent: Towards Interpretable User Modeling via Recurrent Semantics Memory Unit
Zhiqiang Tao (Northeastern University); Sheng Li (University of Georgia); Zhaowen Wang (Adobe Research); Chen Fang (ByteDance AI Lab); Longqi Yang (Cornell University); Handong Zhao (Adobe Research); Yun Fu (Northeastern University)

Interpretable and Steerable Sequence Learning via Prototypes
Yao Ming (Hong Kong University of Science and Technology); Panpan Xu (Bosch Research North America); Huamin Qu (Hong Kong University of Science and Technology); Liu Ren (Bosch Research North America)

Research Track Session RT18: Recommender Systems II – Summit 4, Ground Level, Egan Center
Chair: Xia Ning

Opt: Learn to Regularize Recommender Models in Finer Levels
Yihong Chen (Tsinghua University & Microsoft Research); Bei Chen (Microsoft Research); Xiangnan He (University of Science and Technology of China); Chen Gao (Tsinghua University); Yong Li (Tsinghua University); Jian-Guang Lou (Microsoft Research); Yue Wang (Tsinghua University)

Exact-K Recommendation via Maximal Clique Optimization
Yu Gong (Alibaba Group); Yu Zhu (Alibaba Group); Lu Duan (Zhejiang Cainiao Supply Chain Management Co., Ltd); Qingwen Liu (Alibaba Group); Ziyu Guan (Xidian University); Fei Sun (Alibaba Group); Wenwu Ou (Alibaba Group); Qili Zhu (Shanghai Jiao Tong University)

MeLU: Meta-Learned User Preference Estimator for Cold-Start Recommendation
Hoyeop Lee (NCSOFT Co.); Jinbae Im (NCSOFT Co.); Seongwon Jang (NCSOFT Co.); Hyunsouk Cho (NCSOFT Co.); Sehee Chung (NCSOFT Co.)

DAML: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation
Donghua Liu (Wuhan University); Jing Li (Wuhan University); Bo Du (Wuhan University); Jun Chang (Wuhan University); Rong Gao (Hubei University of Technology)

Enhancing Collaborative Filtering with Generative Augmentation
Qinyong Wang (The University of Queensland); Hongzhi Yin (The University of Queensland); Hao Wang (Alibaba AI Labs); Quoc Viet Hung Nguyen (Griffith University); Zi Huang (The University of Queensland); Lizhen Cui (Shandong University)

OBOE: Collaborative Filtering for AutoML Model Selection
Chengrun Yang (Cornell University); Yuji Akimoto (Cornell University); Dae Won Kim (Cornell University); Madeleine Udell (Cornell University)

1:30PM - 3:30PM

Applied Data Science Track Session ADS9: E-commerce and Advertising – Summit 1, Ground Level, Egan Center
Chair: Anne Kao (Boeing)

SMOILE: A Shopper Marketing Optimization and Inverse Learning Engine
Abhilash Reddy Chenreddy (University of Illinois at Chicago); Parshan Pakiman (University of Illinois at Chicago); Selvaprabu Nadarajah (Information and Decision Sciences); Ranganathan Chandrasekaran (Information and Decision Sciences); Rick Abens (Foresight ROI, Inc.)

Two-Sided Fairness for Repeated Matchings in Two-Sided Markets: A Case Study of a Ride-Hailing Platform
Tom Sühr (Max Planck Institute for Software Systems); Asia J. Biega (Microsoft Research); Meike Zehlike (Max Planck Institute for Software Systems); Krishna P. Gummadi (Max Planck Institute for Software Systems); Abhijnan Chakraborty (Max Planck Institute for Software Systems)

Reserve Price Failure Rate Prediction with Header Bidding in Display Advertising
Achir Kalra (Forbes Media LLC); Chong Wang (S&P Global); Cristian Borcea (New Jersey Institute of Technology); Yi Chen (New Jersey Institute of Technology)

The Identification and Estimation of Direct and Indirect Effects in A/B Tests through Causal Mediation Analysis
Xuan Yin (Etsy, Inc.); Liangjie Hong (Etsy, Inc.)

Personalized Purchase Prediction of Market Baskets with Wasserstein-Based Sequence Matching
Mathias Kraus (ETH Zurich); Stefan Feuerriegel (ETH Zurich)

Research Track Session RT19: Machine Learning Themes II – Summit 2, Ground Level, Egan Center
Chair: Feng Chen

Disambiguation Enabled Linear Discriminant Analysis for Partial Label Dimensionality Reduction
Jing-Han Wu (Southeast University & Ministry of Education); Min-Ling Zhang (Southeast University & Ministry of Education)

ET-Lasso: A New Efficient Tuning of Lasso-type Regularization for High-Dimensional Data
Songshan Yang (Pennsylvania State University); Jiawei Wen (Pennsylvania State University); Xiang Zhan (Pennsylvania State University); Daniel Kifer (Pennsylvania State University)

Isolation Set-Kernel and Its Application to Multi-Instance Learning
Bi-Cun Xu (Nanjing University); Kai Ming Ting (Federation University, Australia); Zhi-Hua Zhou (Nanjing University)

Separated Trust Regions Policy Optimization Method
Luobao Zou (Shanghai Jiao Tong University); Zhiwei Zhuang (Shanghai Jiao Tong University); Yin Cheng (Shanghai Jiao Tong University); Xuechun Wang (Shanghai Jiao Tong University); Weidong Zhang (Shanghai Jiao Tong University)

QuesNet: A Unified Representation for Heterogeneous Test Questions
Yu Yin (University of Science and Technology of China); Qi Liu (University of Science and Technology of China); Zhenya Huang (University of Science and Technology of China); Enhong Chen (University of Science and Technology of China); Wei Tong (University of Science and Technology of China); Shijin Wang (iFLYTEK CO., LTD. & State Key Laboratory of Cognitive Intelligence); Yu Su (iFLYTEK CO., LTD. & State Key Laboratory of Cognitive Intelligence)

Research Track Session RT20: Online and Incremental Algorithms – Summit 3, Ground Level, Egan Center
Chair: Albert Bifet

Scaling Multi-Armed Bandit Algorithms
Edouard Fouché (Karlsruhe Institute of Technology); Junpei Komiyama (University of Tokyo); Klemens Böhm (Karlsruhe Institute of Technology)

Adaptive Deep Models for Incremental Learning: Considering Capacity Scalability and Sustainability
Yang Yang (Nanjing University); Da-Wei Zhou (Nanjing University); De-Chuan Zhan (Nanjing University); Hui Xiong (Rutgers University); Yuan Jiang (Nanjing University)

Streaming Adaptation of Deep Forecasting Models using Adaptive Recurrent Units
Prathamesh Deshpande (IIT Bombay); Sunita Sarawagi (IIT Bombay)

Dual Averaging Method for Online Graph-structured Sparsity
Baojian Zhou (University at Albany, SUNY); Feng Chen (University at Albany, SUNY); Yiming Ying (University at Albany, SUNY)

Factorization Bandits for Online Influence Maximization
Qingyun Wu (University of Virginia); Zhige Li (Shanghai Jiao Tong University); Huazheng Wang (University of Virginia); Wei Chen (Microsoft Research); Hongning Wang (University of Virginia)

AWARDS

SIGKDD Innovation and Service Award

Innovation Award Winner: Charu Aggarwal

Service Award Winner: Balaji Krishnapuram

SIGKDD Test of Time Award

Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, Natalie Glance. Cost-effective Outbreak Detection in Networks. KDD-2007.

KDD 2019 Best Paper Awards

The Best Paper is p1224: Network Density of States by Kun Dong (Cornell), Austin Benson (Cornell), and David Bindel (Cornell).

The Runner Up is p138: Optimizing Impression Counts for Outdoor Advertising by Yipeng Zhang (RMIT University), Yuchen Li (Singapore Management University), Zhifeng Bao (RMIT University), Songsong Mo (Wuhan University), and Ping ZhangHuawei).

Best Applied Track Paper

The Best Paper is p1464: Actions Speak Louder Than Goals: Valuing Player Actions in Soccer by Tom Decroos (KU Leuven), Lotte Bransen (SciSports), Jan Van Haaren (SciSports), Jesse Davis (KU Leuven).

The Runner Up is p594: Developing Measures of Cognitive Impairment in the Real World from Consumer-Grade Multimodal Sensor Streams by Richard Chen (Apple Inc.), Filip Jankovic (Evidation Health, Inc.), Nikki Marinsek (Evidation Health, Inc.), Luca Foschini (Evidation Health, Inc.), Lampros Kourtis (Evidation Health, Inc.), Alessio Signorini (Evidation Health, Inc.), Melissa Pugh (Eli Lilly and Company), Jie Shen (Eli Lilly and Company), Roy Yaari (Eli Lilly and Company), Vera Maljkovic (Eli Lilly and Company), Marc Sunga (Eli Lilly and Company), Han Hee Song (Apple Inc.), Hyun Joon Jung (Apple Inc.), Belle Tseng (Apple Inc.), Andrew Trister (Apple Inc.).

KDD 2019 Dissertation Awards

Award Winner

Tim Althoff, Stanford, advised by Jure Leskovec

Runner Up

Chao Zhang, University of Illinois at Urbana-Champaign, advised by Jiawei Han

Honorable Mention

Michael Yeh, University of California - Riverside, advised by Eamonn Keogh

Honorable Mention

Ioannis (John) Paparrizos, Columbia University, advised by Luis Gravano

KDD 2019 Startup Research Awards:

Participant: Qiang Liu, RealAI

Participant: Chao Liu, TianYanCha

Participant: Zhen Wei, Arkive

Participant: Kartik Yellepeddi, Deepair

KDD 2019 ORGANIZING COMMITTEE

General Chairs
Vipin Kumar, University of Minnesota
Ankur Teredesai, KenSci

Applied Data Science Invited Talk Chair
Deepak Agarwal, LinkedIn
Johannes Gehrke, Microsoft
Sofus Macskássy, HackerRank
Hui Xiong, Baidu / Rutgers University

Applied Data Science Track PC Chairs
Ying Li, Giving Tech Labs
Rómer Rosales, LinkedIn

Best Paper Chairs
Tina Eliassi-Rad, Northeastern University
Jennifer Neville, Purdue University

Deep Learning Day Chairs
Yuxiao Dong, Microsoft Research
Vagelis Papalexakis, UCR

Dissertation Award Chair
Yehuda Koren, Google

Earth Day Chairs
James Hodson, Jožef Stefan Institute
Lucas Joppa, Microsoft
Shashi Shekhar, University of Minnesota

Finance Chair
Ted Senator

Hands-On Tutorial Chairs
Shenghua Bao, Amazon
Wee Hyong Tok, Microsoft

Health Day Chairs
Carly Eckert, KenSci
Faisal Farooq, IBM
Chandan Reddy, Virginia Tech

KDD Cup Chairs
Taposh Dutta-Roy, Kaiser Permanente
Iryna Skrypnyk, Pfizer
Wenjun Zhou, UTK

Local Chairs
Ross Martin, Axiom Data Science
Jessica Austin, Axiom Data Science
Kenrick Mock, University Alaska Anchorage
Lance Ahern, Local Entrepreneur
Brendan Babb, Municipality of Anchorage

Media and Publicity Chairs
Xiang Ren, USC
Derek Young, Amazon
Ping Zhang, The Ohio State University

Plenary Keynote Panel Chairs
Chid Apte, IBM Research
Nitesh Chawla, University of Notre Dame
Geoff Webb, Monash University

Poster Chair
Yong Ge, University of Arizona

Project Showcase Chairs
James Hodson, Jožef Stefan Institute
Dunja Mladenic, Jožef Stefan Institute

Proceedings Chairs
Latifur Khan, University of Texas at Dallas
Jiliang Tang, Michigan State University

Research Track PC Chairs
George Karypis, University of Minnesota
Evimaria Terzi, Boston University

Social Impact Chairs
Vani Mandava, Microsoft Research
Sarah Stone, eScience Institute, UW - Seattle

Sponsorship Chairs
Ritwik Kumar, Netflix
Lei Li, Bytedance AI Lab
Shipeng Yu, LinkedIn

Startup Award Chairs
Sreenivas Gollapudi, Google AI
Xing Xie, Microsoft Research Asia

Student Travel Awards Chair
Juhua Hu, University of Washington, Tacoma

Test of Time Award Chair
Bing Liu, UIC

Tutorial Chairs
Leman Akoglu, CMU
Danai Koutra, University of Michigan

Web and Job Matching Chair
Derek Young, Amazon

Workshop Chairs
Jing Gao, University at Buffalo, SUNY
Anuj Karpatne, Virginia Tech

CONFERENCE**SPONSORS**

SILVERSPONSORS

Capital One · Booz | Allen | Hamilton · lyft · Lambda · we · The Home Depot · NSA (WHERE INTELLIGENCE GOES TO WORK® www.IntelligenceCareers.gov/NSA) · 快手 · inspur

BRONZESPONSORS

ByteDance 字节跳动 · Bloomberg® · Megagon Labs · coinbase · D E Shaw & Co · Square Point · NAVER | LINE · Microsoft · SIEMENS Ingenuity for life · Etsy

KEYNOTESPONSORS

TWO SIGMA · 松鼠AI·智适应 Squirrel AI Learning · GUROBI OPTIMIZATION

RESEARCH TRACKBEST PAPER

Bloomberg® · 平安科技 PING AN TECHNOLOGY

HEALTH DAYBEST PAPER

IBM **Research** AI

WOMEN'S LUNCHEONSPONSOR

Lawrence Livermore National Laboratory

DISSERTATIONAWARD

SIEMENS Ingenuity for life

LANYARDSPONSOR

KenSci

APPLIED DATA SCIENCE TRACKBEST PAPER

Snap Inc. · JD.COM

MEDIASPONSOR

BIG DATA — Mary Ann Liebert, Inc. publishers

BEST STUDENT PAPERAWARD

leidos · TigerGraph · 美团点评 · JD.COM

BAG INSERTSPONSORS

leidos · appen data with a human touch™ · geneia · Microsoft · HexagonML

EXHIBITORS

Cambridge University Press · Springer Nature · frontiers · Lambda · YAHOO! RESEARCH · CRC Press Taylor & Francis Group · indeed · Microsoft · Etsy · TigerGraph · ByteDance 字节跳动 · Square Point · Bloomberg® · GUROBI OPTIMIZATION · 松鼠AI·智适应 Squirrel AI Learning · D E Shaw & Co · neo4j

DIAMONDSPONSORS

PLATINUMSPONSORS

GOLDSPONSORS