Association for Computing Machinery

Advancing Computing as a Science & Profession

San Francisco, CA, USA

KDD 2016

KDD'16

Proceedings of the 22nd ACM SIGKDD International Conference on

Knowledge Discovery and Data Mining

Sponsored by:

ACM SIGKDD and ACM SIGMOD

Association for Computing Machinery

Advancing Computing as a Science & Profession

Notice to Past Authors of ACM-Published Articles

Additional copies may be ordered prepaid from:

Printed in the USA   .

KDD 2016 Chairs' Welcome Message

It is our great pleasure to welcome you to the 2016 ACM Conference on Knowledge Discovery and Data Mining – KDD'16. We hope that the content and the professional network at KDD'16 will help you succeed professionally by enabling you to: identify technology trends early; make new/creative contributions; increase your productivity by using newer/better tools, processes or ways of organizing teams; identify new job opportunities; and hire new team members.

We are living in an exciting time for our profession. On the one hand, we are witnessing the *industrialization of data science*, and the emergence of the *industrial assembly line* processes characterized by the division of labor, integrated processes/pipelines of work, standards, automation, and repeatability. Data science practitioners are organizing themselves in more sophisticated ways, embedding themselves in larger teams in many industry verticals, improving their productivity substantially, and achieving a much larger scale of social impact. On the other hand we are also witnessing astonishing progress from research in algorithms and systems — for example the field of deep neural networks has revolutionized speech recognition, NLP, computer vision, image recognition, etc. By facilitating interaction between practitioners at large companies & startups on the one hand, and the algorithm development researchers including leading academics on the other, KDD'16 fosters technological and entrepreneurial innovation in the area of data science.

This year's conference continues its tradition of being the premier forum for presentation of results in the field of data mining, both in the form of cutting edge research, and in the form of insights from the development and deployment of real world applications. Further, the conference continues with its tradition of a strong tutorial and workshop program on leading edge issues of data mining. The mission of this conference has broadened in recent years even as we placed a significant amount of focus on both the research and applied aspects of data mining. As an example of this broadened focus, this year we have introduced a strong hands-on tutorial program during the conference in which participants will learn how to use practical tools for data mining. KDD'16 also gives researchers and practitioners a unique opportunity to form professional networks, and to share their perspectives with others interested in the various aspects of data mining. For example, we have introduced office hours for budding entrepreneurs from our community to meet leading Venture Capitalists investing in this area. We hope that KDD 2016 conference will serve as a meeting ground for researchers, practitioners, funding agencies, and investors to help create new algorithms and commercial products.

The call for papers attracted a significant number of submissions from countries all over the world. In particular, the research track attracted 784 submissions and the applied data science track attracted 331 submissions. Papers were accepted either as full papers or as posters. The overall acceptance rate either as full papers or posters was less than 20%. For full papers in the research track, the acceptance rate was lower than 10%. This is consistent with the fact that the KDD Conference is a premier conference in data mining and the acceptance rates historically tend to be low. It is noteworthy that the applied data science track received a larger number of submissions compared to previous years. We view this as an encouraging sign that research in data mining is increasingly becoming relevant to industrial applications. All papers were reviewed by at least three program committee members and then discussed by the PC members in a discussion moderated by a meta-reviewer. Borderline papers were thoroughly reviewed by the program chairs before final decisions were made.

In addition, KDD'16 continues to have a strong tutorial and workshop program, as is its tradition since its inception. The program committee reviewed and accepted the following:

Venue or Track	Reviewed	Accepted	Acceptance Rate
Research Track Papers	784	70+72	8.9%+9.2%
Applied Data Science Track Papers	331	40+26	12%+7.9%
Workshops	47	17	36.2%
Tutorials	37	11	29.7%
Applied Data Science Talks	Invited	12	Invited
Hands-on Tutorials	Invited	8	Invited
Regular Keynotes	Invited	4	Invited
Turing Award Winner Talk	Special	1	Special

In the table above the research track papers and applied data science track papers are broken up into the full presentations and poster presentations. We have four excellent keynote presentations from Jennifer Chayes, Joe Hellerstein, Greg Papadopoulous, and Nando de Freitas. In addition, we are honored to have a Turing Award winner, Whitfield Diffie, as one of the keynote speakers. We also encourage attendees to attend the keynote and invited talk presentations.

Putting together *KDD'16* was a team effort. We first thank the authors for providing the content of the program. We are grateful to the program committee and the senior program committee, who worked very hard in reviewing papers and providing feedback for authors. Finally, we thank the numerous sponsors and ACM SIGKDD for their support in hosting the conference.

We hope that you will find this program interesting and thought-provoking and that the conference will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world.

Charu Aggarwal
Alex Smola
KDD'16 Research Track Chairs

Rajeev Rastogi
Dou Shen
KDD'16 Applied Data Science Track Chairs

Balaji Krishnapuram
Mohak Shah
KDD'16 General Chairs

Table of Contents

Keynote Talks

Panel

Applied Data Science Track Oral Papers

Applied Data Science Track Invited Talks

Applied Data Science Track Posters

Research Track Oral Papers

Research Track Posters

Tutorials

KDD 2016 Conference Organization

General Chairs:	Balaji Krishnapuram *(IBM)*
	Mohak Shah *(Bosch)*
Research Track Program Chairs:	Alex Smola *(Amazon)*
	Charu Aggarwal *(IBM)*
Applied Data Science Track Program Chairs:	Dou Shen *(Baidu)*
	Rajeev Rastogi *(Amazon)*
Associate General Chair	Shipeng Yu (LinkedIn)
Workshop Chairs:	Quoc Le *(Google)*
	Zhi-Hua Zhou *(Nanjing University)*
Tutorial Chairs:	Hanghang Tong *(Arizona State University)*
	S V N Vishwanathan *(UCSC)*
KDD Cup Chairs:	Markus Weimer *(Microsoft)*
	Shou-De Lin *(National Taiwan University)*
	Vijay Narayanan *(Microsoft)*
Panel Chair:	Andrei Broder *(Google)*
Applied Data Science Invited Talks Chairs:	Evangelos Simoudis *(Synapse Partners)*
	Usama Fayyad *(Barclays Bank and Oasis500)*
	Rajesh Parekh *(Facebook)*
Poster Chair:	Jiliang Tang *(Yahoo Labs)*
Best Paper Chairs:	Jiliang Tang *(Yahoo Labs)*
Doctoral Dissertation Award Chair:	Haixun Wang *(Facebook)*
Innovation and Service Award Chair:	Jian Pei *(Simon Fraser University)*
Test-of-Time Award Chair:	Wei Wang *(UCLA)*
Local Arrangements Chairs:	Joaquin Quiñonero Candela *(Facebook)*
	Olivier Chapelle *(Google)*
Student Travel Award Chairs:	Jure Leskovec *(Stanford University)*
	Leman Akoglu *(Stony Brook University)*
Media and Publicity Chairs:	Ankur Teredesai *(University of Washington)*
	Gabor Melli *(Opengov)*
Proceedings Chairs:	Rómer Rosales *(LinkedIn)*
	Stefanie Jegelka *(MIT)*

Social Networking Chairs:	Faisal Farooq *(IBM)*
	Ron Bekkerman *(University of Haifa)*
Sponsorship Chair:	Claudia Perlich *(Dstillery)*
	Yu Zheng *(Microsoft Research)*
Treasurer:	Ying Li *(EV Analysis)*
Video and Streaming Chairs:	Jie Tang *(Tsinghua University)*
	Marko Grobelnik *(Jozef Stefan Institute)*
Web Chairs:	David Hazel *(University of Washington)*
	Derek Young *(Seasonal View)*
Job Match Chair:	Unmesh Kurup *(Bosch)*
Social Impact Chair:	Rayid Ghani *(University of Chicago)*
Research Track Senior Program Committee:	Alice Oh *(KAIST)*
	Ambuj Singh *(UCSB)*
	Amr Ahmed *(Google)*
	Andrew Tomkins *(Google Research)*
	Anil Vullikanti *(Virginia Tech)*
	Aristides Gionis *(Aalto University)*
	Arno Siebes
	Bart Goethals *(University of Antwerp)*
	Chih-Jen Lin *(National Taiwan University)*
	Christos Faloutsos *(Carnegie Mellon University)*
	Corinna Cortes *(Google)*
	Dacheng Tao *(The University of Technology, Sydney)*
	Deepak Agarwal *(LinkedIn)*
	Diane Cook *(Washington State University)*
	Dimitrios Gunopulos
	Eamonn Keogh *(UC Riverside)*
	Edo Liberty *(Yahoo Labs)*
	Evgeniy Gabrilovich *(Google)*
	Evimaria Terzi *(Boston University)*
	Fei Wang *(University of Connecticut)*
	Francesco Bonchi *(Yahoo Barcelona)*
	George Karypis
	Graham Cormode *(Warwick)*
	Guojun Qi *(University of Central Florida)*
	Hong Cheng *(Chinese University of Hong Kong)*
	Huan Liu *(ASU)*
	Hui Xiong *(Rutgers)*
	Ian Davidson *(University of California-Davis)*
	James Bailey

Applied Data Science Track Senior Program Committee:

Anitha Kannan *(Facebook)*
Balaraman Ravindran *(Indian Institute of Technology Madras)*
Brian D'Alessandro *(Facebook)*
Chao Liu *(Jindi Data Technology Ltd Co.)*
Chiranjib Bhattacharya *(Indian Institute of Science Bangalore)*
Divesh Srivastava *(AT&T Research)*
Gautam Das *(University of Texas Arlington)*
Gautam Shroff *(Tata Consultancy Services Research)*
Graham Williams *(Australian Taxation Office)*
Haixun Wang *(Google)*
Hang Li *(Huawei Technologies)*
Hema Raghavan *(LinkedIn)*
Hui Xiong *(Rutgers University)*
Jan Pedersen *(Microsoft)*
Jeffrey Pan *(Facebook)*
Jun Yan *(Microsoft Research Asia)*
Minos Garofalakis *(University of Chania)*
Ping Luo *(Chinese Academy of Sciences)*
Ralf Herbrich *(Amazon)*
Rayid Ghani *(University of Chicago)*
Roberto Bayardo *(Google)*
S. Muthukrishnan *(Rutgers Univ)*
Sinno Pan *(Nanyang Technological University)*
Tina Eliassi-Rad *(Rutgers Univ)*
Vincent Tseng *(National Chiao Tung University)*
Vipin Kumar *(University of Minnesota)*
Wayne Thompson *(SAS)*
Wei Fan *(Baidu)*
Xiaoming Jin *(Tsinghua University)*
Ying Li *(EV Analysis Corp)*
Yoelle Maarek *(Yahoo! Labs)*
Yu Zheng *(Microsoft Research)*

KDD 2016 Sponsors & Supporters

Sponsor: ACM SIGKDD

Supporters: Diamond

Supporters: Platinum

Supporters: Gold

Orchestrating a brighter world

NEC

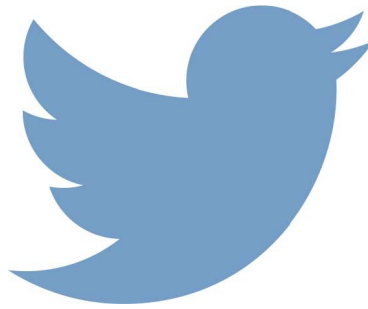**Supporters: Silver**

ebay

Tencent 腾讯

SIEMENS

Ingenuity for life

UBER

sas

criteo labs

Supporters: Bronze

+ableau

NETFLIX

Google

Booking.com

Supporters: Lanyard Sponsor

Supporters: KDD CUP Sponsor

Supporters: Student Travel Awards

Supporters: Program Ad Sponsor

Supporters: Media Sponsor

BIG DATA

Mary Ann Liebert, Inc. publishers

Exhibitors:

INDIANA UNIVERSITY
DATA SCIENCE PROGRAM

MORGAN & CLAYPOOL
PUBLISHERS

NON SOLUS

ELSEVIER

Springer

Georgia Institute of Technology