



16th Australasian Data Mining Conference (AusDM 2018)

Bathurst, Australia,
28–30 November 2018



<http://ausdm18.ausdm.org/>

The Australasian Data Mining Conference has established itself as the premier Australasian meeting for both practitioners and researchers in data mining. It is devoted to the art and science of intelligent analysis of (usually big) data sets for meaningful (and previously unknown) insights. This conference will enable the sharing and learning of research and progress in the local context and new breakthroughs in data mining algorithms and their applications across all industries.

Specifically, the conference seeks to showcase: Research Prototypes; Industry Case Studies; Practical Analytics Technology; and Research Student Projects. AusDM18 will be a meeting place for pushing forward the frontiers of data mining in academia and industry. AusDM 2018 will offer a variety of keynote speakers and tutorials on important and interesting contemporary topics in data mining.



Since AusDM02 the conference has showcased research in data mining, providing a forum for presenting and discussing the latest research and developments. Built on this tradition, AusDM18 will facilitate the cross-disciplinary exchange of ideas, experience and potential research directions. This year AusDM is proud to announce that conference proceedings will be published in Springer's Communication in Computer and Information Science. Previous proceedings (from 2006 to 2016), have been printed as volumes in the CRPIT series.

Publication and topics

We are calling for papers, both research and applications, and from both academia and industry, for presentation at the conference. All papers will go through double-blind, peer-review by a panel of international experts. Accepted papers will be published in the AusDM 2018 proceedings by Springer. Some selected papers will be invited for submission with extension in a special edition of a Springer journal. Please note that we require that at least one author for each accepted paper will register for the conference and present their work. One full registration will cover at most two papers.

AusDM invites contributions addressing current research in data mining and knowledge discovery as well as experiences, novel applications and future challenges. Topics of interest include, but are not restricted to:

- Applications of Data Mining and Case Studies
- Big Data Analytics
- Biomedical and Health Data Mining
- Business Analytics
- Computational Aspects of Data Mining
- Data Integration, Matching and Linkage

- Data Mining Education
- Data Mining in Security and Surveillance
- Data Preparation, Cleaning and Preprocessing
- Data Stream Mining
- Implementations of Data Mining in Industry
- Integrating Domain Knowledge
- Knowledge Discovery and Presentation
- Link, Tree, Graph, Network and Process Mining
- Multimedia Data Mining
- Mobile Data Mining
- New Data Mining Algorithms
- Privacy-preserving Data Mining
- Spatial and Temporal Data Mining
- Text Mining
- Web and Social Network Mining

Keynote speakers

As is tradition for AusDM we have lined up an excellent keynote speaker program. Each speaker is a well known researcher and/or practitioner in data mining and related disciplines. The keynote program provides an opportunity to hear from some of the world's leaders on what the technology offers and where it is heading.

Professor Junbin Gao, The University of Sydney

Junbin Gao is Professor of Big Data Analytics at the University of Sydney Business School. Until recently his major research interest has been machine learning and its application in data science, image analysis, pattern recognition, Bayesian learning & inference, and numerical optimization etc. He is the author of 260 academic research papers and two books. His recent research has involved new machine learning algorithms for big data in business. He won two research grants in Discovery Project theme from the prestigious Australian Research Council.

Professor Geoff Webb, Monash University

Geoff Webb is Director of the Monash University Centre for Data Science. He is a leading data scientist and the recipient of multiple Australian and International awards. He developed many of the key mechanisms of support-confidence association discovery in the 1980s. His OPUS search algorithm remains the state-of-the-art in rule search. He pioneered multiple research areas as diverse as black-box user modelling, interactive data analytics and statistically-sound pattern discovery, and his algorithms are widely deployed.

Submission of papers

We invite three types of submissions for AusDM 2018:

- **Academic submissions:** Regular academic submissions can be made in Research Track reporting on research progress, with a paper length up to 12 pages. For academic submissions we will use a double-blind review process, i.e. paper submissions must NOT include author names or affiliations (and also not acknowledgements referring to funding bodies). Self-citing references should also be removed from the submitted papers (they can be added on after the review) for the double blind reviewing purpose.
- **Industry submissions:** Submissions can be made in the Application Track to report on specific data mining implementations and experiences in governments and industry projects. Submissions in this category can be up to 12 pages. The review process for these submissions will also be double-blind. A special committee made of industry representatives will assess industry submissions.
- **Industry Showcase submissions:** Submission from industry and government on an analytics solution that has raised profits, reduced costs and/or achieved other important policy and/or business outcomes can be made in this track with a one page Abstract only. The review process for these submissions will also be double-blind.

Paper submissions are required to follow the general format specified for papers. LaTeX styles and Word templates will be available while LaTeX will be the recommended typesetting package.

The electronic submissions must be in PDF only, and made through the AusDM'18 Submission Page.

Special Tracks

In addition to our regular academic track AusDM 2018 invites papers in the following special categories. Papers accepted from each of these tracks will be presented together in a session. These special tracks are:

- Image Data Mining
- Identification Through Data Mining

- Mobile and Sensor Network Data Mining Track
- Statistics in Data Science

Papers submitted in these tracks will be reviewed in the same process as those submitted to the regular academic track. All accepted papers will be published in the same proceedings.

Image Data Mining

One picture is worth more than thousands of words. Image data potentially contains a wealth of information and widely exists in diverse fields. Therefore, developing data mining technologies to extract useful patterns and knowledge from image data is of great importance and significance. In the current age of big data, digital images and videos are flying through cyberspace. Large-scale image/video analysis and security/privacy have been attracting an increasing attention. But due to the unstructured nature of image data, analyzing and interpreting it for object detection and recognition and behaviour analysis from the multi-discipline perspective remains a challenging problem in the field of computer vision and data mining.

This special track aims to provide a platform for researchers to discuss and showcase the distinctive theoretical concepts, new mining models and cutting-edge mining technology related to image and video data.

Potential topics include, but are not limited to:

- Image Feature extraction and classification
- Object detection and recognition from images/videos
- Image retrieval and knowledge discovery
- Information hiding, digital watermarking and steganography
- Recent machine learning applications in image data mining

Identification Through Data Mining

Identification is a fundamental task in various disciplines such as computer vision, cybersecurity, digital forensics, and biometrics. A substantial amount of research efforts has been devoted to developing identification approaches applied in the above-mentioned fields. Meanwhile, the data generated in different knowledge areas has explosively increased in the past few decades. The ever-increasing amount and variety of data open new opportunities in boosting the performance of identification, but the large-scale size of data also presents new challenges for identification systems. Utilizing data mining techniques to gain useful insights from massive data becomes crucial for many identification tasks.

The aim of this special track is to showcase the recent advances in identification systems driven by data mining techniques. An important focus is on identification methods capable of working with large-scale data acquired in camera surveillance network and social media. We solicit high-quality original research papers that advance the development of identification systems through data mining techniques. Submitted papers should not be previously published or be under consideration for publication elsewhere.

Potential topics include, but are not limited to:

- Biometric identification on large-scale databases
- Human characteristics prediction and inference from social media data
- People identification/re-identification in camera surveillance networks
- Anomaly detection and identification in camera surveillance networks
- Intrusion detection and identification in computer networks
- Provenance-oriented identification/clustering in large-scale databases and social media
- Preference identification for recommending purpose

Mobile and Sensor Network Data Mining

There has recently been a considerable amount of research work on using data compression techniques to minimise the volume of transmitted traffic, and consequently assist in reducing power consumption levels in Wireless Sensor Networks. Data management and processing for wireless sensor networks have become a topic of active research in several fields of computer science, such as the distributed systems, the database systems, and the data mining. The main aim of deploying the WSNs-based applications is to make the real-time decision which has been proved to be very challenging due to the highly resource-constrained computing, communicating capacities, and the huge volume of fast-changed data generated by WSNs. This challenge motivates the research community to explore novel data mining techniques dealing with extracting knowledge from large and continuously arriving data from WSNs. Traditional data mining techniques are not directly applicable to WSNs due to the nature of sensor data, their special characteristics, and limitations of the WSNs.

This track aims to showcase the recent advances in data mining techniques in mobile and sensor networks. We solicit high-quality original research papers that advance the development of identification systems through data mining techniques. Submitted papers

should not be previously published or under consideration for publication elsewhere.

Potential topics include, but are not limited to:

- Data mining techniques in mobile and sensor networks
- Data stream processing in mobile and sensor networks
- Data fusion techniques in mobile and sensor networks
- Social data mining through distributed mobile sensing
- Application of data mining techniques in mobile and sensor networks
- Distributed data mining techniques for mobile and sensor networks
- Mobile and sensor network data analytics
- Challenges for data mining in distributed mobile and sensor networks

Statistics in Data Science

Intensive-data-driven research is empowering theoretical breakthroughs and high-tech innovations, enabling new methodologies in academic discovery, and offering new sustainable means to solve significant societal and economic challenges and understand the world. Data science is a very fast growing research domain which embedded the combination of computational (i.e. computer-intensive) and inferential (i.e. statistics-oriented) thinking, with the notion of Microdata to Big Data. Since the core theories in computer science and statistical science were developed separately, there is an oil and water problem to be surmounted in data science.

For an example, the basic statistical theory does not have a place for runtime and other computational resources while core computer science theory does not have a place for statistical risk and inferential resources. As well, the most appealing challenge in Microdata is the simulation of detailed population characteristics, while in Big Data is the potential of personalised attributes. This session welcomes the full range of submissions, ranging from methodological contributions to case studies involving statistics based modellings or decisions making process (e.g., Bayesian thinking, spatial statistics, machine learning methods, modern statistics, or novel modelling techniques) in data science.

We ask that each submission include a statement about possible implications for resolving some computational and inferential challenges in data science with special focus to data mining, combining data, computation, and inferences that is as wide as from computing statistics or running machine learning algorithms to estimating reliability measures including standard errors and confidence intervals on their outputs in any fields.

Important Dates

Paper Submission Closed:	Friday 20 July 2018
Authors Notified:	Monday 1 October 2018
Camera Ready Submission:	Monday 15 October 2018
Preliminary Program Available:	Wednesday 31 October 2018
Early Bird Cut-Off Date for Authors:	Friday 2 November 2018
Conference Dates:	Wednesday 28 – Friday 30 November 2018

Organising Committee

Conference Chairs	Zahid Islam, Charles Sturt University Chang-Tsun Li, Charles Sturt University
Program Chairs (Academic)	Md. Rafiqul Islam, Charles Sturt University Yun Sing Koh, University of Auckland
Program Chairs (Industry)	Yanchang Zhao, CSIRO, Sydney David Stirling, University of Wollongong
Program Chair (Industry Showcase)	Warwick Graco, Australian Taxation Office
Proceedings Chair	Kok-Leong Ong, La Trobe University
Tutorial Chairs	Jixue Liu, University of South Australia Yee Ling Boo, RMIT University
Web Chairs	Michael Furner, Charles Sturt University Ling Chen, University of Technology Sydney
Competition Chairs	Tony Nolan, Australian Taxation Office Paul Kennedy, University of Technology Sydney
Publicity Chairs	Ji Zhang, University of Southern Queensland Ashad Kabir, Charles Sturt University
Track Chairs (Image Data Mining)	Lihong Zheng, Charles Sturt University

Xufeng Lin, Charles Sturt University

Track Chairs (Sensor Data Mining)

Quazi Mamun, Charles Sturt University
Sabih Rehman, Charles Sturt University

Track Chairs
(Statistics in Data Science)

Azizur Rahman, Charles Sturt University
Ryan Ip, Charles Sturt University
Hien Nguyen, La Trobe University, Melbourne

Track Chairs
(Identification through Data Mining)

Xufeng Lin, Charles Sturt University
Xingjie Wei, University of Bath, UK

Steering Committee Chairs

Simeon Simoff, University of Western Sydney
Graham Williams, Microsoft

Steering Committee

Peter Christen, The Australian National University, Canberra
Ling Chen, University of Technology, Sydney
Zahid Islam, Charles Sturt University
Paul Kennedy, University of Technology, Sydney
Jiuyong (John) Li, University of South Australia, Adelaide
Richi Nayak, Queensland University of Technology, Brisbane
Kok-Leong Ong, La Trobe University, Melbourne
Dharmendra Sharma, University of Canberra
Glenn Stone, Western Sydney University
Andrew Stranieri, Federation University Australia, Mount Helen
Yanchang Zhao, CSIRO, Sydney

Join us on LinkedIn: <http://www.linkedin.com/groups/AusDM-4907891>