



PMDB 2022 @ DASFAA 2022

PMDB 2022

**1st Workshop on Pattern Mining and Machine Learning
in Big Complex Databases**

in conjunction with DASFAA 2022



11th April 2022

Motivation

- **Machine learning (ML)** has become a key research area, with highly effective techniques
- But many machine learning techniques operate as black boxes.
- **Pattern mining (PM)** aims at developing algorithms to find patterns in data meeting user-defined constraints.
- PM algorithms may not be as effective as ML for prediction tasks but can uncover patterns that are easily interpretable.

Motivation

- **Key issue with PM and ML research** is the focus on simple data types
- **Real life:** complex data types such as temporal data, streaming data, spatial data, multimedia data, text data, heterogeneous data, and graph data
- Thus, we need to adapt ML and PM models for **complex data**.
- **Another key issue:** scaling current ML and PM techniques to handle **big complex data**

Organization committee

[Philippe Fournier-Viger](#)

Shenzhen University



[Mourad Nouioua](#)

Harbin Institute of Technology :



[Hamido Fujita](#)

Iwate Prefectural University



[Lin Zhang](#)

Tencent.com



[Vincent S. Tseng](#)

National Chiao Tung University



Program

- 1 keynote speaker
- 10 submitted papers
- Double-blind review
- 6 accepted papers (60%)
- The proceedings of **PMDB** will be jointly published with other workshops after the conference, by Springer.

Keynote speaker:

Distinguished Prof. Joshua Zhexue Huang
(Shenzhen University, China)

"Approximate Computing for Big Data Analysis"



Research papers

9:45 Paper #2 An Algorithm for Mining Fixed-Length High Utility Itemsets ([PDF](#))

Le Wang

10:05 Paper #3 A Novel Method to Create Synthetic Samples with Autoencoder Multi-layer Extreme Learning Machine ([PDF](#))

Qihang Huang, Yulin He, Shengsheng Xu and Joshua Zhexue Huang

10:25 Paper #4 Pattern Mining: Current Challenges and Opportunities ([PDF](#))

Philippe Fournier-Viger, Wensheng Gan, Youxi Wu, Mourad Nouioua, Wei Song, Tin Truong and Hai Duong Van

10:45 Paper #7 Why not to Trust Big Data: Identifying Existence of Simpson's Paradox

Rahul Sharma, Minakshi Kaushik, Sijo Arakkal Peious, Mahtab Shahin, Ankit Vidhyarthi and Dirk Draheim

11:05 Paper #8 Localized Metric Learning for Large Multi-Class Extremely Imbalanced Face Database ([PDF](#))

Seba Susan and Ashu Kaushik

11:25 Paper #9 Top-k dominating queries on incremental datasets

Jimmy Ming-Tai Wu, Ke Wang and Jerry Chun-Wei Lin



Best paper award

THIS ACKNOWLEDGES THAT THE PAPER
**Localized Metric Learning for Large Multi-Class
Extremely Imbalanced Face Database**
by
Seba Susan and Ashu Kaushik

HAS BEEN SELECTED AS THE BEST PAPER AT THE 1ST WORKSHOP ON PATTERN MINING
AND MACHINE LEARNING IN BIG COMPLEX DATABASES (**PMDB 2022**)
HELD ONLINE AT THE DASFAA 2022 CONFERENCE

A handwritten signature in black ink that reads "Philippe".

Philippe Fournier-Viger
on behalf of the PMDB 2022 committee

11th April 2021

Date

Next year

- Try to expand this workshop to a larger audience.
- A potential journal special issue