World Congress 2021
www.worldcongressdsa.com

The Frontiers in Intelligent Data and Signal Analysis

New York, USA
July 11 - 22, 2021

17th International Conference on Machine Learning and Data Mining
MLDM 2021

July 17 – 22, 2021
www.mldm.de
Preface

The fourteenth event of the International Conference on Machine Learning and Data Mining MLDM 2018 was held in New York (www.mldm.de) running under the umbrella of the World Congress "The Frontiers in Intelligent Data and Signal Analysis, DSA2017" (www.worldcongressdsa.com).

After the peer-review process, we accepted 86 high-quality papers for oral presentation. The topics range from theoretical topics for classification, clustering, association rule and pattern mining to specific data-mining methods for the different multimedia data types such as image mining, text mining, video mining, and Web mining. Extended versions of selected papers will appear in the international journal Transactions on Machine Learning and Data Mining (www.ibai-publishing.org/journal/mldm).

The tutorial days rounded up the high quality of the conference. Researchers and practitioners got an excellent insight in the research and technology of the respective fields, the new trends and the open research problems that we like to study further.

A tutorial on Data Mining, a tutorial on Case-Based Reasoning, a tutorial on Intelligent Image Interpretation and Computer Vision in Medicine, Biotechnology, Chemistry and Food Industry, and a tutorial on Standardization in Immunofluorescence were held before the conference.

We would like to thank all reviewers for their highly professional work and their effort in reviewing the papers. We would also thank the members of the FutureLab Artificial Intelligence IBaI-2 (www.futurelab-ai-ibai-2.de), who handled the conference as secretariat. We appreciate the help and understanding of the editorial staff at ibai-publishing house, who supported the publication of these proceedings (http://www.ibai-publishing.org/html/proceeding.php).
Last, but not least, we wish to thank all the speakers and participants who contributed to the success of the conference. See you in 2021 in New York at the next World Congress (www.worldcongressdsa.com) on “The Frontiers in Intelligent Data and Signal Analysis, DSA2022”, which combines under its roof the following three events: International Conferences on Machine Learning and Data Mining, MLDM (www.mldm.de), the Industrial Conference on Data Mining, ICDM (www.data-mining-forum.de), and the International Conference on Mass Data Analysis of Signals and Images in Medicine, Biometry, Drug Discovery Biotechnology, Chemistry and Food Industry, MDA.

July 2021

Petra Perner
18th International Conference on Machine Learning and Data Mining – MLDM 2021

www.mldm.de

**Program Chair**

Petra Perner ..................................... IBaI Leipzig, Germany

**Program Committee**

Michelangelo Ceci  
Univestiy of Bari, Italy

Sung-Hyuk Cha  
Pace University, USA

Ireneusz Czarnowski  
Gdynia Maritime University, Poland

Christoph F. Eick  
Univestiy of Houston, USA

Mark J. Embrechts  
Rensselaer Polytechnic Institute and CardioMag Imaging, Inc, USA

Ana Fred  
Technical University of Lisboa, Portugal

Giorgio Giacinto  
University of Cagliari, Italy

Robert Haralick  
City University of New York, USA

Aminata Kane  
Concordia University, Canada

Olga Krasotkina  
Russian Stae University, Russia

Dimitris Karras  
Chalkis Institute of Technology, Greece

Adam Krzyzak  
Concordia University, Canada

Valerio Pascucci  
University of Utah, USA

Dan Simovici  
University of Massachusetts Boston, USA

Francis E.H. Tay  
National University of Singapore, Singapore

Patrick Wang  
Northeastern University, USA

Zeev Volkovich  
ORT Braude College of Engineering, Israel

Patrick Wang  
Northeastern University, USA
Instructions for Presentations

Dear Author,

Please come in the morning. Due to the corona situation will freely chance the schedule.

Please note: Presentation time for long paper will be 20 Minutes. 15 Minutes for the talk and 5 minutes for discussion.

A beamer will be available. Please bring your presentation on stick. **Please make sure that you will be available the whole conference. We will freely change the schedule since we have to expect that some people will not get their visa for whatever reasons. We are working hard to support those people but sometimes you cannot do anything. That means we will put available people to the free slots.**

The sessions do not have a thematic arrangement. We have found out over time that this is more preferable since running one session on the same subject can get sometimes a bit exhaustive. People also like to get to know about subjects they are not working on so that they get a good overview about the topic of the conference.

For the poster presenters. You will put your poster on the wall of the room- Sticks are available. Please do it when you arrive and leave the poster on the wall until you leave.

Each poster presenter should be prepared for a 5 minutes summarization of his work. I will step with the audience from poster to poster. After all posters have been visited you can start with a personal discussion of interested people.

The poster should be 90x90 cm.

I wish us an inspiring and fruitful conference.

Best regards,

Prof. Dr. Petra Perner
Chair of MLDM
Publication 2021

Please note the publication can be download from http://www.ibai-publishing.org/html/proceeding.php. It is an OPEN ACCESS publication. Everyone in the world can freely download the publication. Besides that, the proceedings book will appear on researchgate and EU agenda.

Reference is:

Proceedings, Machine Learning and Data Mining in Pattern Recognition, Petra Perner (Ed.), 17th International Conference on Machine Learning and Data Mining, MLDM 2021, New York, USA, July 17-22, 2021, ibai-publishing, ISSN (Print) 1864-9734, ISSN (Online) 2699-5220
MLDM 2021 – Program

Monday, July 19, 2021 – First Conference Day

8:00 - 9:00 am Registration

9:00 - 9:15 am Welcome
   *Petra Perner*

1st Session
   *Session Chair:*

09:15-09:35 The Generalization of Mathematical Morphology to Non-numeric Sets
   Xiaojin Ye¹², Robert Haralick¹
   ¹ The Graduate Center, CUNY, New York, USA,
   ² Fordham University, Bronx, New York, USA

09:35-09:55 AdeNet: Deep learning architecture that identifies damaged electrical insulators in power lines
   Ademola Okerinde, Lior Shamir, William Hsu, Tom Theis
   Kansas State University, 2164 Engineering Hall, Manhattan, USA

09:55-10:15 Generalized Approach to Locality-Sensitive Hashing and Learned Aggregation
   Arun Luthra, Edmond Abrahamian², Jean Luo²,
   Ana Armenta², and Prince Paulraj³
   ² AT&T Chief Data Office, Palo Alto, USA,
   ³ AT&T Chief Data Office, Plano, USA

10:15-10:30 Break

2nd Session
   *Session Chair:*

10:30-10:50 Virtual Patient Data Generation Using Wasserstein Generative Adversarial Networks for Randomized Clinical Trial Data Sharing
   Daihao Xue¹, Andre Couturier², Chathuri
10:50-11:10 Learning to Predict with Unavailable Features: an End-to-End Approach via Knowledge Transferring Chun-Chen Lin, Li-Wei Chang, Chun-Pai Yang, and Shou-De Lin National Taiwan University, Taipei, Taiwan

11:10-11:30 Semantic Text Matching using Transfer Learning and Deep Learning Nagendra Kumar¹, Jung-jae Kim², Shubhi Bansal³
¹ Indian Institute of Technology Indore, India, ² Institute for Infocomm Research, A*STAR, Singapore, ³ Indian Institute of Technology Indore, India

11:30-11:45 Break

3rd Session
Session Chair:

11:45-12:05 Link Prediction on Graphs with Variational Graph Stacked Autoencoders Maria Vaida Harrisburg University of Science and Technology, Harrisburg, USA

12:05-12:25 On the Nature of Negative Sampling: How Non-Accident Data Help Us Understand Accident Occurrence Peter Way, Jeremiah Roland, Mina Sartipi Department of Computer Science and Engineering, University of Tennessee at Chattanooga, Chattanooga, USA

12:25-12:45 Application of Divisive Clustering for Reducing Bias in Imbalanced Data Diandra Prioleau, Kiana Alikhademi, Armisha Roberts, Joshua Peeples, Alina Zare, Juan E. Gilbert University of Florida, College of Engineering,
Gainesville FL

12:45-14:00 Lunch Time

4th Session
Session Chair:
14:00-14:20 Does really a CNN recognize at the object?
Yuya Kase and Masaomi Kimura
Shibaura Institute of Technology, Tokyo, Japan

14:20-14:40 Domain adaptation with unlabeled data for model transferability between airborne particle identifiers
Predrag Matavulj¹, Sanja Brdar¹, Miloš Racković², Branko Šikoparija³, Ioannis N. Athanasiadis³
¹ BioSense Institute, University of Novi Sad, Serbia,
² Faculty of Sciences, University of Novi Sad, Serbia,
³ Geo-Information and Remote Sensing Group, Wageningen University and Research, The Netherlands

14:40-15:00 Evaluating Preferential Voting Methods for Multiple Class Classification
Vaibhav Kumar Katturu¹, ParamPuneet Kaur Thind¹, Teryn Cha², Sung-Hyuk Cha¹
¹ Computer Science Department, Pace University New York, NY, USA, ² Mathematics, Engineering Technologies & Computer Sciences, Essex County College, Newark, NJ, USA

15:00-15:15 Break

5th Session
Session Chair:
15:15-15:35 Differentiating Between Covid19 and Viral Pneumonia, using Deep Learning Models Ensemble Methods
Ben Zagagy, Maya Herman, Ofer Levi
The Open University of Israel, Haifa, Israel

15:35-15:55 Symbiotic Hybrid Neural Network Watchdog for Outlier Detection
Justin Bui, Robert J. Marks II
Baylor University, Waco, TX, 76798, USA
15:55-16:15  Monitoring Social Media Content for Religious Context
            Manika Mittal, Shivangi Goel, V. Pratishtha Sharma, Sonia Khetarpaul
            Department of Computer Science and Engineering, Shiv Nadar University, Delhi NCR India

16:15-16:30  Break

16:30-16:50  Cross-Domain Recommender System based on Neural Collaborative Filtering
            Nawaf Alharbi and Doina Caragea
            Kansas State University, Manhattan, USA

16:50-17:10  Bonsai: A Framework for Convolutional Neural Network Acceleration Using Criterion-Based Pruning
            Joseph Bingham¹,² and Sam Helmich²
            ¹ Rutgers University, New Brunswick, USA
            ² John Deere Intelligent Solutions Group, Urbandale, USA

17:10-17:20  Fairwell

17:20  End of Conference
MLDM 2021 – Tutorials
http://www.mldm.de/tutorials.php

Saturday, July 17, 2021
9:00 am – 5:00 pm  Case-Based Reasoning
                  Prof. Dr. Petra Perner

Sunday, July 18, 2021
9:00 am – 5:00 pm  Data Mining
                  Prof. Dr. Petra Perner

Thursday, July 13, 2021
9:00 am – 3:00 pm  Intelligent Image Processing
                  Prof. Dr. Petra Perner
Announcement for Publication in the Special Issue of the Intern. Journal Transaction on Machine Learning and Data Mining
(http://www.ibai-publishing.org/journal/mldm/about.php)

Dear Authors,

You are cordially invited to submit a revised version of your paper to the Intern. Journal Transaction on Machine Learning and Data Mining of the October 2021 Issue. Researchers, who are interested to submit their paper, should submit it until September 1st, 2021.

Papers who arrive later will not be considered for the October issue. They can be published if accepted in the July 2022 issue.

I am looking forward to your excellent contributions.

Best regards,
Prof. Dr. Petra Perner
Editor of Trans on MLDM
Announcement 2022

Dear Participants and Authors,

We hope to see you again in 2022. Please come and join us for the 18th International Conference on Machine Learning and Data Mining MLDM 2022 (www.mldm.de).
For the latest news please visit www.mldm.de or go to www.worldcongressdsa.com.

I am looking forward to see you all again.
Thank you for choosing MLDM 2021.

Sincerely yours,

Prof. Dr. Petra Perner
Chair of MLDM and Head of the Machine Learning and Data Mining Community
A Note on the Citation Policy or How to strengthen up the Conference and Your Own Work!

Dear Author,

to ensure that the proceedings and a particular paper from the proceedings is properly cited, we would highly recommend that you follow the citation style given below.
If you do not properly follow the citation style given below you might risk that your paper or any of the other papers is not properly recognized by the citation index.
Thank you for taking care of this important issue.

Kind regards,
Prof. Dr. Petra Perner
Chair of the Conference

Recommended Citation Style of the Proceedings and the Papers


A particular paper:

Example for Poster Proceedings:

Example for Workshop Proceedings:

Example for Journal:
A Note on Citation Policy or How to strengthen up the Conference and Your Own Work!

As you might know, the value of a publication, a conference or a workshop is judged based on the citation rate. American publications and conferences are highly ranked based on this statistical measure. American publishers make sure that people accept this policy and publish papers that cite the respective publications published in the journal or at the conference where the researchers submit their paper.

If you want your publication to be considered a valuable one, you should follow this policy when you publish at the conferences ICDM, MLDM or MDA. The papers of ICDM, MLDM and MDA are published in the Springer LNCS or LNAI Series and by ibai-publishing Prof. Dr. Petra Perner and most of the libraries worldwide have a copy of these proceedings. Also, the library index at http://www.informatik.uni-trier.de/~ley/db/ carefully lists all publications that have been published at these conferences.

All OPEN ACCESS journals of ibai-publishing (http://www.ibai-publishing.org/journal/index.html) are also listed there:

• International Journal Transactions on Machine Learning and Data Mining
• International Journal Transactions on Case-Based Reasoning
• International Journal Transactions on Mass Data Analysis of Signals and Images.

Any researcher can access this index for free. It shows title, authors and their affiliation, and abstract. It provides a researcher with a fast and complete overview of all publications presented at a conference or a workshop or published in a journal.

So please, before you submit a paper, check the index for relevant publications. If you have only cited publications from outside the main stream of ICDM, MLDM or MDA or any of the journals, the question will be raised why you then did not publish your paper outside the main stream.

If the reviewer suggests to cite a publication, please do so!
It will make your publication more valuable.
Journals by ibai-publishing

The journals are free on-line journals but having in parallel hardcopies of the journals. The free on-line access to the content of the paper should ensure fast and easy access to new research developments for researchers all over the world. The hardcopy of the journal can be purchased by individuals, companies, and libraries.

Transactions on Machine Learning and Data Mining
(ISSN: 1865-6781)

The International Journal “Transactions on Machine Learning and Data Mining” is a periodical appearing twice a year. The journal focuses on novel theoretical work for particular topics in Data Mining and applications on Data Mining.

Net Price (per issue): EURO 100
Germany (per issue): EURO 107 (incl. 7% VAT)

Submission for the journal should be send to:
info@ibai-publishing.org

For more information visited: www.ibai-publishing.org/journal/mldm/about.html

Transactions on Case-Based Reasoning
(ISSN:1867-366X)

The International Journal “Transactionson Case-Based Reasoning” is a periodical appearing once a year.

Net Price (per issue): EURO 100
Germany (per issue): EURO 107 (incl. 7% VAT)

Submission for the journal should be send to:
info@ibai-publishing.org

For more information visited: www.ibai-publishing.org/journal/cbr/about.html
Transactions on Mass-Data Analysis of Images and Signals
(ISSN:1868-6451)

The International Journal "Transactions on Mass-Data Analysis of Images and Signals" is a periodical appearing once a year.
The automatic analysis of images and signals in medicine, biotechnology, and chemistry is a challenging and demanding field. Signal-producing procedures by microscopes, spectrometers and other sensors have found their way into wide fields of medicine, biotechnology, economy and environmental analysis. With this arises the problem of the automatic mass analysis of signal information. Signal-interpreting systems which generate automatically the desired target statements from the signals are therefore of compelling necessity. The continuation of mass analyses on the basis of the classical procedures leads to investments of proportions that are not feasible. New procedures and system architectures are therefore required.

Net Price (per issue): EURO 100
Germany (per issue): EURO 107 (incl. 7% VAT)

Submission for the journal should be send to:
info@ibai-publishing.org

For more information visited: www.ibai-publishing.org/journal/massdata/about.php
ibai-publishing headed by Prof. Dr. Petra Perner (IAPR Fellow) publishes your scientific work in high quality with an ISBN-number or an ISSN-number.

Why choose ibai-publishing?
- Highly recognized scientists in the field will review your work which will serve as a basis for publication.
- Announcement of books, journals and proceedings will be disseminated to a large scientific community worldwide.
- Your work will be reviewed and published at earliest convenience.

Costs
Publication costs of books depend on individual factors (such as size of book, number of pages, number of books, softcover or hardcover, data format etc.). On request we will send you a quotation. Please contact us at: info@ibai-publishing.org

Advertisement
- International E-marketing, more than 50,000 scientists and institutions around the world will get the advertisement of your publication.
- Sale and distribution through ibai-publishing Prof. Dr. Petra Perner.

Further Information
ibai-publishing
Prof. Dr. Petra Perner
Fabrikstraße 10
04637 Falkenberg
Germany

Phone: +49 (0) 30 469 99 35 44 • FAX: +49 (0) 30 469 99 35 46
E-mail: info@ibai-publishing.org • Web: www.ibai-publishing.org
Newark in New Jersey

Dear Participant,

We are running the conference in Newark. That is the little sister of NY and located in the State of New Jersey. It used to be a business area with a lot of companies producing goods like textiles, iron and other stuff. The raw material came from upstate New Jersey and was transported by ship on the Passaic River. Money was made fast here and you can still see the signs of this past. There are old business houses at Main Street, as you might know them from old movies, very nicely renovated and preserved. The Town House has a golden cupola. I still remember the time when one of my relatives, who had been a weaver and later on in America a textile-company owner, proudly said: "We have a golden cupola now."

Americans are proud of what they have and of what they have achieved. They are active members of the community. They do not play down their performances, something other nations could learn from them.

Unfortunately, they do not rely so much on old things. If it has made his job, it goes to ruin. You can still see the landing bridges of the ships that brought the raw material to Newark. The technique must have been an innovation at the time when it was installed. The landing bridges are located near the Main Railway Station, in an area where more or less corrosion on the steel bridges and cranes has done its job. A bit of color on the industrial facilities would make a nice open industrial museum of this area.

The Main Railway Station is another point of interest worthwhile looking at. Go to the waiting hall, have a look around, visit the pub
there. Stroll around downtown Newark and get a flavor of the atmosphere of life there.

Not very long ago it has not been so easy to walk around in Newark. The unemployment rate was high and mainly black people were seen on the streets. Newark was an area where living was inexpensive. The government then installed a lot of education and training programs and created new jobs. The result is evident. People are proud now, whether they are white, black, or yellow. They have undergone education or training and hold certificates. They have new jobs now that are in accordance with the special needs of the new economy.

The officer at the Immigration Office wanted to know what I am doing and I answered that I am a scientist and working in something like artificial intelligence. He looked at me, congratulated me for the job I am doing, but then he said: “Artificial intelligences destroys jobs. What shall people do who are not so smart? What jobs can they get?” And I told him that we want to make life easier and more beautiful, that to that end we needed AI, E-Commerce etc. I pointed out that we also know that we had to create new jobs for those who could not be scientists and the new economy that we were creating needed jobs for people at all levels of intelligence and that we had to educate them for that.

Further, see the Parliament of Newark which looks a little like Washington’s Parliament.

Share a taxi and go to Hoboken. Walk around the main street with its small shops and restaurants. You can take from Hoboken Main Railway Station the bus to Point Authority. The bus brings you to NY within minutes. The railway station in Hoboken is marvelous and a must see. It is in Tiffany style and gives you a taste of the past. From there you can take a ferry to
different places in NY. You can go to Battery Park, to the Finance District and other points of interest.

Seeing NY is a must anyway for everybody who has not been there yet. From Battery Park the underground runs below 5th Avenue and brings you to all the important places – to museums, galleries, parks, shopping areas and restaurants. I especially like the City Museum of NY. It is a small but very beautiful and informative multimedia installation about the beginning of NY and how the city has developed over time.

The Metropolitan Museum is another highlight, but it takes a lot of time to see everything. Before you go there, make sure what time period or what area you want to visit and concentrate on that. Next time you will be in NY visit the Metropolitan again and decide for another time period. So, step by step you will see everything.

Greenwich Village is a place where you can find nice restaurants offering different styles of cuisine. Plan to stop at lunch or dinner time in Greenwich Village, relax over there in a garden restaurant and enjoy the food.

There is a bus going to Newark airport that can bring you back to Newark. You can catch the bus from different points along the main street in Manhattan. Another option is to go to Penn Station and take the fast train to Newark.

Enjoy the Conference and Have Fun by Visiting Newark, Hoboken and New York!

Sincerely yours,
Prof. Dr. Petra Perner
Chair

July 2021