

KDD 2020 Call for Research Papers

Key Dates

- Submission: February 13, 2020
- Notification: May 15, 2020
- Camera-ready: June 1, 2020
- Short Promotional Video (Required): July 1, 2020
- Conference (San Diego, California): August 22-27, 2020

Description

We invite submission of papers describing innovative research on all aspects of knowledge discovery and data mining, ranging from theoretical foundations to novel models and algorithms for data mining problems in science, business, medicine, and engineering. Visionary papers on new and emerging topics are also welcome, as are application-oriented papers that make innovative technical contributions to research. Authors are explicitly discouraged from submitting incremental results that do not provide major advances over existing approaches.

All deadlines are at 11:59PM [Alofi Time](#). There will be absolutely no exception to these deadlines.

Topics of interest include, but are not limited to:

- Data Science: Methods for analyzing scientific and business data, social networks, time series; mining sequences, streams, text, web, graphs, rules, patterns, logs data, IoT data, spatio-temporal data, biological data; recommender systems, computational advertising, multimedia, finance, bioinformatics.
- Big Data: Large-scale systems for text and graph analysis, machine learning, optimization, sampling, parallel and distributed data mining (cloud, map-reduce, federated learning), novel algorithmic and statistical techniques for big data.
- Foundations: Models and algorithms, asymptotic analysis; model selection, dimensionality reduction, relational/structured learning, matrix and tensor methods, probabilistic and statistical methods; deep learning, meta learning, AutoML, reinforcement learning; classification, clustering, regression, semi-supervised and unsupervised learning; personalization, security and privacy, visualization; fairness, interpretability and robustness.

Submission Directions

KDD is a dual track conference hosting both a Research track and an Applied Data Science track. Due to the large number of submissions, papers submitted to the Research track will not be considered for publication in the Applied Data Science track and vice versa. Authors are encouraged to read the track descriptions carefully and to choose an appropriate track for their submissions. Submissions are limited to a total of nine (9) pages, including all content and

references, and must be in PDF format and formatted according to the new Standard ACM Conference Proceedings Template. For LaTeX users: unzip acmart.zip, make, and use sample-sigconf.tex as a template;

Additional information about formatting and style files is available online at: <https://www.acm.org/publications/proceedings-template>.

Papers that do not meet the formatting requirements will be rejected without review. In addition, authors can provide an optional two (2) page supplement at the end of their submitted paper (it needs to be in the same PDF file and start at page 10) focused on reproducibility. This supplement can only be used to include (i) information necessary for reproducing the experimental results, insights, or conclusions reported in the paper (e.g., various algorithmic and model parameters and configurations, hyper-parameter search spaces, details related to dataset filtering and train/test splits, software versions, detailed hardware configuration, etc.), and (ii) any pseudo-code, or proofs that due to space limitations, could not be included in the main nine-page manuscript, but that help in reproducibility (see reproducibility policy below for more details).

The Research track follows a double-blind review process. Submitted papers must not include author names and affiliations and they must be written in a way so that they do not break the double-blind reviewing process. If the preliminary version of a paper was posted in arXiv, the authors should NOT mention as their own paper in the submission. Papers that violate the double-blind review requirements will be desk rejected.

Website for submissions: <https://easychair.org/conferences/?conf=kdd2020>.

Important Policies

Reproducibility

Submitted papers will be assessed based on their novelty, technical quality, potential impact, insightfulness, depth, clarity, and reproducibility. Authors are strongly encouraged to make their code and data publicly available whenever possible. In addition, authors are strongly encouraged to also report, whenever possible, results for their methods on publicly available datasets. Algorithms and resources used in a paper should be described as completely as possible to allow reproducibility. This includes experimental methodology, empirical evaluations, and results. The authors are encouraged to take advantage of the optional two-page supplement to provide the appropriate information. The reproducibility factor will play an important role in the assessment of each submission.

Authorship

Every person named as the author of a paper must have contributed substantially both to the work described in the paper and to the writing of the paper. Every listed author must take responsibility for the entire content of a paper. Persons who do not meet these requirements may

be acknowledged, but should not be listed as authors. Post-submission changes to the author list are not allowed.

Dual Submissions

No dual submissions are allowed. Submitted papers must describe work that is substantively different from work that has already been published, or accepted for publication, or submitted in parallel to other conferences or journals.

Violations on the dual submission policy may lead to immediate desk rejection and further penalties including prohibition of submitting to conferences and journals sponsored by SIGKDD or/and ACM for a certain period. The employers of the violating authors may be notified.

Conflicts of Interest

During the submission process, enter the email domains of all institutions with which you have an institutional conflict of interest. You have an institutional conflict of interest if you are currently employed or have been employed at this institution in the past three years, or you have extensively collaborated with this institution within the past three years. Authors are also required to identify all PC/SPC members with whom they have a conflict of interest, e.g., advisor, student, colleague, or coauthor in the last five years.

Additional information about ACM's Conflict of Interest policy, which KDD follows, can be found at <https://www.acm.org/publications/policies/conflict-of-interest>.

Retraction Policy

KDD follows ACM's policies, which are described at <https://www.acm.org/publications/policies/retraction-policy>.

Attendance

For each accepted paper, at least one author must attend the conference and present the paper. Authors of all accepted papers must prepare a final version for publication, a poster, and a three-minute short video presentation (details will be in the acceptance notification).

Copyright

Accepted papers will be published in the conference proceedings by ACM and also appear in the ACM Digital Library. The rights retained by authors who transfer copyright to ACM can be found here.

AUTHORS TAKE NOTE: The official publication date is the date the proceedings are made available in the ACM Digital Library. This date for KDD 2020 is on or after July 15, 2020. The official publication date affects the deadline for any patent filings related to published work.

Contact Information

Email: research2020@kdd.org

Wei Wang and Heng Huang

Research Track PC co-Chairs of KDD-2020

Potential Additions:

Fairness and Trust: Algorithmic fairness, trustworthy systems. FAIR data.

Data Science in Public Policy: use, risks, and benefits of data science for a well-defined vertical application area: such as democratic practice, media, trade in digital services, health, education, and societal infrastructure for energy, communication, transportation, and national security. Advances in theory, methods and tools that help us understand the opportunities and challenges that digital data pose to markets, organizations, society, and government; methods and evaluating policies to advance these opportunities and address these challenges, while considering both technical constraints and constraints related to social, political and economic feasibility.

Digital and Data Infrastructure: software systems and software-hardware integrated systems that serve as platforms to enable and amplify the impact of data science algorithms and applications. Systems for sensing and sourcing data, systems for storing, querying, learning, and analytics over data, streaming methods for analyzing very large datasets and systems for securely deploying data science methods in high-impact application verticals.

In foundational areas include: statistical foundations, high-dimensional data, large-scale hypothesis testing, regularization and sparsity, causal inference, false discovery minimization methods.