Kanav Mehra

kanav.mehra@uwaterloo.ca | Homepage: kanav-mehra.github.io | LinkedIn: kanav-mehra | Github: kanav-mehra

Summary: Experienced data scientist adept in designing and implementing production machine learning systems with a graduate research background in responsible AI and natural language processing (NLP).

EXPERIENCE

Beatdapp | *Data Scientist* — Sep 2023 - Present
- Implemented an end-to-end production ML system to identify fraud in music streaming and royalty claims leveraging ensemble graph community detection algorithms, centrality features, and tree-based methods.
- Developed an unsupervised graph representation learning framework in Python and Neo4j to learn user embeddings from music streaming behaviour, improving precision in downstream fraud detection tasks by 4%.

University of Waterloo | *Graduate AI Researcher* — Aug 2021 - Aug 2023
- Developed a fair-ranking algorithm for search engine and recommendation systems that guarantees fairness across multiple protected attributes and beats state-of-the-art on aggregate fairness and relevance benchmarks by 21%.
- Designed fair deliberation and voting methods using synthetically generated election data to enhance user welfare and representation outcomes by 18%, improving minority user inclusion by 65% in collective decision-making.

PwC | *Software Consultant* — Jul 2019 - Sep 2021
- Led a 3-member global team to implement an application monitoring system with real-time anomaly detection, customized alerting profiles, and automated incident generation. Reduced incident detection time by 50%.
- Managed the development of a threat response system involving two global teams. Investigated log data and implemented a spam prediction model using Splunk Machine Learning Toolkit, reducing threats to users by 73%.

Data Science for Social Good Foundation | *Data Scientist* — Jul 2020 - Nov 2020
- Leveraged large-scale Twitter discourse (500k tweets in 3 languages) to extract narratives and identify unmet needs during natural disasters, aiding relief efforts with actionable insights.
- Designed a pipeline using zero-shot text classification with a large fine-tuned BART model for topic analysis, followed by sentiment analysis, point-of-view extraction, and summarization to flag tweets indicating unmet needs.
- Utilized document embedding model to generate tweet and user embeddings to construct a network followed by community detection and network clustering methods to extract popular users and narratives.

Indian Institute of Technology, Kharagpur | *Research Intern* — May 2017 - Jul 2017
- Developed novel unsupervised and supervised ensemble text summarization algorithms using graph processing, community detection, and learning-to-rank methods, beating base methods by 9%.

EDUCATION

University of Waterloo — Waterloo, Canada
Master of Mathematics, Computer Science (Thesis) — 2021 - 2023
Advisor: Prof. Kate Larson | Cumulative GPA: 96.33%
Thesis: Fairness and Diversity in Ranking and Voting Systems

Indian Institute of Engineering Science and Technology, Shibpur — West Bengal, India
Bachelor of Technology in Information Technology — 2015 - 2019
Distinction: First Class Honours

TECHNICAL SKILLS

Programming Languages: Python, C, SQL, Splunk SPL, PromQL
Technologies: PyTorch, Sklearn, Git, NetworkX, Google Cloud, Vertex AI, MLFlow, Neo4j, BigQuery, Elasticsearch
Key Skills: Machine Learning, Predictive Modelling, NLP, Model Development and Validation, Graph Learning

SELECTED PUBLICATIONS

Mehra, K., Sreenivas, N., Larson, K. (2023). *Deliberation and Voting in Approval-Based Multi-Winner Elections.* International Joint Conference on Artificial Intelligence (IJCAI) 2023 Main Track.

Crayton, A., Fonseca, J., **Mehra, K.**, Ng, M., Ross, J., Sandoval-Castañeda, M. & von Gnechten, R. (2020). *Narratives and Needs: Analyzing Experiences of Cyclone Amphan Using Twitter Discourse.* Tackling Climate Change with Machine Learning Workshop at Neural Information Processing Systems (NeurIPS) 2020.

Dutta, S., Chandra, V., **Mehra, K.**, Das, A., Chakraborty, T. & Ghosh, S. (2018). *Ensemble Algorithms for Microblog Summarization.* IEEE Intelligent Systems, Issue on "Summarization of Things", vol. 33, no. 3, pp. 4–14.

HONORS & AWARDS

Best Paper Award at Citizen-Centric Multiagent Systems Workshop, AAMAS	2023
Pradeep Khare Memorial Scholarship	2023
University of Waterloo Entrance Scholarship	2021
International Master's Award of Excellence	2021
PwC STAR&R Client Appreciation Award	2020
GAABESU Undergraduate Research Award for excellence in research	2019