Karan Vombatkere

kvombat@bu.edu • *kvombatkere* • *cs-people.bu.edu/kvombat*

Professional Summary

Applied ML researcher and data scientist with experience developing models in research and production environments. Strong background in optimization, NLP, graph-based modeling, and end-to-end ML pipelines to support data-driven decision-making.

Technical Skills: Python [PyTorch, Pandas, scikit-learn], SQL, Machine learning [supervised/unsupervised learning, NLP, graph neural networks, optimization], Generative AI fundamentals [transformer architectures, prompting principles, LLM evaluation concepts], REST APIs, AWS, Spark, Git

Education

Boston University — Boston, MA
Ph.D. Computer Science — *Aug 2021 - 2026 (Expected)*
- *Algorithmic Data Mining, Machine Learning, Social Computing* [GPA: 3.90]

University of Rochester — Rochester, NY
M.S. Data Science — *May 2018*
- *Computational & Statistical Methods* [GPA: 3.83]

B.S. Electrical & Computer Engineering — *May 2017*
B.A. Physics
- *Highest Distinction, Magna Cum Laude* [GPA: 3.92]

Applied AI & Research Experience

Boston University, Boston, MA — *Aug 2021 - present*
Ph.D. Researcher — Advisor: Dr. Evimaria Terzi

Approximation methods for balancing submodular utility and cost
- Designed approximation algorithms with guarantees and graph-based ML methods for solving submodular utility-cost tradeoffs in team formation, recommender systems, and influence maximization.
- Implemented models in Python and evaluated performance against greedy and integer-programming baselines on large real-world datasets to quantify approximation and runtime tradeoffs.

Max Planck Institute-SWS, Germany — *May - Aug 2022*
Research Intern — Advisor: Dr. Krishna Gummadi

Content personalization in social media feeds
- Designed analytical framework to audit personalization of social media recommendation algorithms.
- Applied statistical methods to evaluate exploration/exploitation dynamics on 5M videos from TikTok; found that 30–50% of recommended videos reflected inferred user interests, with engagement signals (watch time, skips, likes, follows) significantly driving personalization levels.
- Collaborated with interdisciplinary research teams on algorithmic transparency and model evaluation.

IBM, Cambridge, MA — *Sep 2018 - Jun 2021*
Data Scientist — Supervisor: Dr. Mark Freeman

Machine Learning for Bid Price Optimization
- Designed and deployed hierarchical decision-tree and logistic regression models in Python for large-scale B2B bid pricing optimization, trained on millions of historical bids.
- Built end-to-end ML pipeline including feature engineering and model validation, and developed a production REST API microservice delivering real-time pricing decisions at 2ms latency for Verizon.

Natural Language Processing for Clinical Authorization
- Developed an NLP rules engine in Python to extract structured clinical features from unstructured patient records; deployed the framework on AWS for CVS Health.
- Enabled AI-assisted surgical pre-authorization workflows, contributing to a 5–10× reduction in review timelines by supporting clinician decision-making.

Large-Scale Data Engineering
- Built SparkSQL pipelines to process millions of enterprise records, enabling real-time data updates for executive dashboards at Apple Media Products.

Brand Networks, Rochester, NY *Jan - May 2018*
Data Engineer Intern
- Identified optimal Facebook ad-campaign configurations using SQL scripts. Developed classification models in Python to predict KPIs and presented a metric-driven campaign configuration process.

Audio Research Lab, University of Rochester, NY *May - Aug 2016*
Xerox Research Fellow Mentor: Dr. Zhiyao Duan
- Developed an automated lyric display system for live music performances in Java. Used a real-time implementation of the dynamic time warping algorithm to align annotated and live temporal sequences.

Selected Publications

Vombatkere, Terzi. **Computing Approximate Pareto Frontiers for Submodular Utility and Cost Tradeoffs**. Preprint 2026.

Vombatkere, Lappas, & Terzi. **A QUBO Framework for Team Formation**. European Conference on Machine Learning and Principles of Knowledge Discovery in Databases 2025.

Vombatkere, Gionis, & Terzi. **Forming Coordinated Teams that Balance Task Coverage and Expert Workload**. Springer Data Mining and Knowledge Discovery 2025.

Vombatkere, Mousavi, Zannettou, Roesner, & Gummadi. **TikTok and the Art of Personalization: Investigating Exploration & Exploitation on Social Media Feeds**. The Web Conference 2024.

Vombatkere, Terzi. **Balancing Task Coverage and Expert Workload in Team Formation**. SIAM International Conference on Data Mining 2023.

Kritharakis, Luo, Unnikrishnan, & Vombatkere. **Detecting Trends in Streaming Financial Data using Apache Flink**. ACM International Conference on Distributed & Event-Based Systems 2022.

Teaching Experience

Boston University - *Teaching Fellow* *2022 - 2025*
- Assisted undergraduate and graduate courses in algorithmic data mining, database systems, linear algebra, and combinatorics; designed labs and problem sets and mentored students.

University of Rochester - *Teaching Assistant* *2014 – 2018*
- Supported instruction across engineering, mathematics, and physics courses; facilitated problem-solving workshops, recitations and labs.

Honors and Awards

Boston University Teaching Excellence Award
University of Rochester Merit Scholarships *[Genesee Scholarship, Dean's Scholarship]*
Awarded full scholarship for undergraduate tenure.
Citation for Achievement in College Leadership
Awarded for demonstrating outstanding undergraduate teaching and research commitment.
Donald M. Barnard Engineering Prize
Awarded annually to one senior for high personal achievement in Electrical & Computer Engineering.
Tau Beta Pi Engineering Honor Society *[National Tau Beta Pi Scholarship]*
Phi Beta Kappa Honor Society