Brian Hentschel

bhentschel9@gmail.com | https://bhentsch.github.io/

Education

Harvard University May 2022
Doctor of Philosophy in Computer Science

Harvard University May 2017
Master of Science in Computer Science

Pomona College May 2014
B.A Mathematics, B.A Computer Science

Professional Experience

Senior Research Scientist - Information Retrieval Current Position
Pinecone

- Applied research role focusing on semantic search (finding relevant data in unstructured data like text)
- Current research on training relevance models (text embedding and reranking models)
- Led algorithmic development of proprietary approximate nearest neighbor search algorithms at Pinecone - work allowed Pinecone to go from pod-based to serverless architecture
- Created novel quantization schemes for high dimensional data improving recall from 95% to 99% at same memory

Ph.D. Student 2017-2022
Harvard University *Cambridge, MA*

- Thesis on using machine learning to tailor data structures and algorithms to data properties
- Authored multiple patented data structures used in creation of ML Systems Startup
- Given awards for 1) best paper at a conference and 2) a top paper in field
- Published 10 papers (5 first author)

Research Intern 2018
Microsoft Research: Data Mining, Management, and Exploration Team

- Project focuses on how to use machine learning to automatically manage databases in the cloud
- Used reinforcement learning and cost estimation on database query plans to suggest index configurations in Azure SQL Server AutoAdmin

Software Engineer 2014-2015
LinkedIn Publishing Team

- Developed data pipelines, middle tier and backend logic for "who viewed my posts"

Selected Awards

SIGMOD Research Highlight Award (top 5 paper in data systems for the year)
Best Paper, *International Conference on Extending Database Technology* (EDBT)
Harvard Smith Graduate Fellowship - given to exceptional PhD students in engineering
Harvard Distinction in Teaching Award (2X)

Selected Publications

Exact PPS Sampling with Bounded Sample Size **Information Processing Letters**
Brian Hentschel, Peter J. Haas, and Yuanyuan Tian

Entropy-Learned Hashing **SIGMOD 2022**
Brian Hentschel, Utku Sirin, and Stratos Idreos.

MotherNets: Rapid Deep Ensemble Learning **SysML 2020**
Abdul Wasay, Brian Hentschel, Yuze Liao, Sanyuan Chen, and Stratos Idreos.

General Temporally-Biased Sampling for Online Model Management **TODS 2019**
Brian Hentschel, Peter J. Haas, and Yuanyuan Tian