Blake Bullwinkel

bbullwinkel@microsoft.com • blakebullwinkel.com • Seattle, WA • (917) 328-1377

Education

Harvard University — Cambridge, MA
M.S. in Data Science. GPA 3.95/4 — *May 2022*

Williams College — Williamstown, MA
B.A. in Mathematics, Chinese. GPA 3.83/4 (*cum laude*) — *June 2020*

University of Oxford — Oxford, UK
Attended as part of the selective, year-long Williams-Exeter Program at Oxford (WEPO). — *June 2019*

Professional Experience

Microsoft — Redmond, WA
Data & Applied Scientist — *Aug 2022–Present*
• Deployed an LLM-powered app that answers questions about internal documentation using retrieval augmented generation (LangChain, Azure OpenAI Service, Azure Cognitive Search).
• Created a vector database of text embeddings to correlate performance bugs and identify common root causes.
• Built a pipeline to detect kernel-mode memory leaks across the Azure fleet and collaborated with OS engineers to mitigate the highest impact leaks, saving around 2TB of memory per day.
• Trained ML models that help deployment teams assess the risk of Azure Host OS updates.

Harvard University — Cambridge, MA
Teaching Fellow — *Feb–May 2022*
• Selected to assist professors in teaching of CS 109b: Advanced Topics in Data Science, a course focused on non-linear statistical methods and deep learning models, including CNNs, RNNs, LSTMs, GANs, and transformers.

PepsiCo R&D — Valhalla, NY
Data Science & Analytics Intern — *May–Aug 2021*
• Developed a Python package for anomaly detection of water usage time series data.
• Trained models (ARIMA, LSTM, FB Prophet) to forecast the water usage efficiency of 17 beverage plants.
• Developed an automated data pipeline with actionable insights in Power BI, adopted nationwide.

Research Experience

Transfer Learning with Physics-Informed Neural Networks — *Feb 2022–Dec 2022*
• Developed a multi-head architecture and transfer learning procedure for efficient simulation of branched flows.
• Paper accepted to the *Machine Learning and the Physical Sciences* workshop at NeurIPS 2022.

Generative Adversarial Network Methods for Solving Differential Equations — *Feb 2021–May 2022*
• Researched methods to improve the training stability of DEQGAN, a generative adversarial network for solving differential equations, developed novel transfer learning algorithms, and implemented methods in a PyTorch package.
• Paper accepted to the *AI for Science* workshop at ICML 2022.

The Fairness Impact of Differentially Private Synthetic Data — *Sept 2021–May 2022*
• Worked with Microsoft data scientists to research the fairness impact of differentially private synthetic data in ML.
• Paper accepted to the *Theory and Practice of Differential Privacy* workshop at ICML 2022.

Honors and Awards

Certificate of Distinction in Teaching for CS109b based on student ratings (mean 4.67/5) — *2022*
IACS Student Scholarship to support data science thesis research ($20,000 award) — *2021*
Goldberg Prize in Mathematics for the best mathematics colloquium (department-wide senior prize) — *2020*
Linen Prize in Chinese for achieving distinction in Chinese (department-wide senior prize) — *2020*
Carolyn Korthals Altes Scholarship for academics and potential to contribute to society — *2019*

Skills and Interests

Programming — Python (NumPy, pandas, sklearn, TensorFlow, PyTorch), R, SQL, HTML/CSS, JavaScript
Tools/Platforms — Conda, Jupyter, Git, Docker, Kubernetes, Azure, AWS
Language — Working proficiency in written and spoken Chinese (Mandarin)
Interests — Running, rowing, piano, writing (Medium blog), Rubik's cube solving (WCA profile)