

VEDANT GAUR

Philadelphia, PA // San Francisco, CA

☎ 650-762-9475 ✉ vedantg@wharton.upenn.edu 🌐 vedantgaur.com 📄 [vedantgaur](https://www.linkedin.com/company/vedantgaur) 🎧 [vedantgaur](https://www.youtube.com/channel/UCv33333333333333333333)

Education

University of Pennsylvania, Jerome Fisher M&T Program

B.S.E., Computer Science (Engineering) and B.S. Economics, Finance (Wharton)

May 2027

Philadelphia, PA

University of Pennsylvania

M.S.E., Computer Science

Expected 2027

Philadelphia, PA

Technical Skills

Languages: Python, Java, C/C++, SQL, JavaScript, R, OCaml, \LaTeX

ML/DS: PyTorch, TensorFlow, NumPy, Pandas **NLP/RAG:** LangChain, vector DBs

Platforms: AWS, GCP, Ubuntu **Data:** PostgreSQL, MongoDB **Tools:** GitHub

Selected Publications

Gaur, V., Goyal, T., Liu, L. "When should we (not) trust automatic LLM benchmarks?" **EMNLP 2025** (First-author)

Gaur, V., Saunshi, N. "Reasoning in large language models through symbolic math word problems," **ACL 2023** (First-author)

4 additional publications (*NeurIPS ATTRIB 2024, IEEE URTC 2022*) — Full list: vedantgaur.com/publications

Experience

Amazon

June 2025 – August 2025

SDE Intern

Seattle, WA

- Engineered and deployed ML pipelines for intent classification, sentiment analysis, and routing prediction on scrubbed, multi-channel customer transcripts, for transfer reduction and contact routing improvement opportunities.
- Integrated RAG and LLM-based task classification to support natural language querying and actionable insight generation.
- Established automated data preprocessing, model retraining, and monitoring workflows with GDPR-compliant storage (Gluon S3), reducing manual analysis time by 80% and increasing actionable contact reduction from 5% to a projected 30%.

Stanford NLP Group

July 2024 – May 2025

Researcher

Stanford, CA

- Researching multi-trigger classification in mechanistic interpretability under Dr. Jan-Philipp Fränken to improve transparency in neural networks.
- Led model performance evaluations, open-sourced Anthropic's paper and met single-classification accuracy AUROC of $\sim 98\%$.

Princeton PLI

June 2022 – May 2025

Lead Author & Researcher

Princeton, NJ

- First-author EMNLP 2025 on LLM benchmarking with Dr. Tanya Goyal and Dr. Leqi Liu, improving cross-model reliability by $\sim 20\%$.
- First-author ACL 2023 on symbolic math reasoning with Dr. Nikunj Saunshi, attaining $\sim 80\%$ accuracy on benchmark tasks.
- Built robust evaluation protocols and analyses for symbolic reasoning datasets to reduce evaluation variance.

Wharton Analytics Fellows

September 2023 – Present

Senior Analyst

Philadelphia, PA

- [IKEA] Led development of an LLM to enhance internal dashboard navigation, optimizing user experience and streamlining data access across departments.
- [Zillow] Modeled marketing spend for ZHL submissions, improving efficiency by 10%. Presented to Zillow's C-Suite and contributed to internal marketing mixture model package refinement.
- [Fox] Developed clustering algorithms for viewer personas, increasing targeted viewership by 15%.

BOLD AI

November 2021 – January 2023

Founder & CEO

- Led clinical testing in Bay Area clinics for a low-cost Transformer-based ($F1 \sim 0.87$) tool for detecting stutters in speech, working with 20+ SLPs and academics.

Projects

Generative Engine Optimization Dashboard | Real-time insights about site visibility + key target metrics

June 2025

RAG-Based Research Literature Review Tool | Automated retrieval and summarization | *Pytorch, Pinecone*

June 2024

Deep Learning Live Interview Platform | Live coding + model evaluation | *Python, React, MERN Stack, AWS*

June 2023

Honors & Hackathons

HackMIT (2x winner), PennApps, CalHacks, YHack, xAI, AGI House

Leadership & Involvement

M&T Student Board (Treasurer + Alumni Relations Student Representative), Wharton Analytics Fellows (Senior Analyst), Wharton Asia Exchange, MUSE Consulting, Wharton Undergraduate Entrepreneurship Club, Venture Initiation Program