

Veer Pareek

+1 (650) 740-9177 | vpareek2@illinois.edu | vpareek2.github.io

Education

University of Illinois Urbana-Champaign

Bachelor of Sciences in Data Science, Minors in Computer Science, French

Champaign, IL

Aug. 2021 – May 2025

Experience

Machine Learning Intern

June 2023 – August 2023

Striim

Palo Alto, CA

- Harnessed Generative AI to build a model that translates natural language queries into actionable, compliant TQL (SQL copilot) code.
- Built a consumer-facing documentation chatbot using OpenAI embeddings and LangChain, with Google search functionality. First intern project to be incorporated in the distributed platform.

Data Science Intern

May 2022 – August 2022

NextPhase.AI

San Mateo, CA

- Enhanced data automation quality through visualization using Apache Superset; developed parameterized dashboards in Power BI and Excel from employee skills survey data, enabling managers to filter by skills like Technical, People, and Leadership.
- Integrated ZoomInfo to help prospect higher quality deals. Used profiling criteria such as technology stacks and current platforms used, along with categories like revenue, location, etc., to curate a list of targets.

Projects

AlphaGAT / C, CUDA

May 2024 – Present

- Reimplementation of Google's original AlphaZero paper in pure C and CUDA for maximum efficiency
- Enhanced the algorithm by replacing the ResNet with a GAT (Graph Attention Network) to improve the representation of the chessboard in the Monte Carlo Tree Search algorithm
- Still in progress, but as of now achieving faster training times with similar chess ELO, working on improving it

CereLex / Python, PyTorch, CUDA

June 2024 – Present

- Wake word detection using EEGs to attempt to replicate a "Hey Siri/ Alexa" from EEG scans using the ZuCo2 dataset
- Implemented a streaming transformer in torch for accuracy chunk detection, rather than full sentence
- Training in CUDA for speedup

BitDB / Python, C, Ruby, MLX

November 2023 – January 2024

- SQLite Database Clone written in one 1600-line C file, offers CRUD operations and more
- Built an assistant named Ada for natural language database queries for the specific database language (simpler SQL)
- Powered by a LoRA fine-tuned Mistral 7B model implemented and tuned in Apple's MLX

Skills

Languages: C/C++, Python, Java, SQL, HTML, CSS, R

Libraries: PyTorch, CUDA, TensorFlow, JAX, TRLX

Frameworks: Node.js, Flask, Swift, MySQL, Neo4j, MongoDB, Redis

Concepts: Deep Learning, High-Performance Computing, Research and Development, Data Analysis, CI/CD