Khush Gupta

(972) 832-5760 | khushgx@gmail.com | linkedin.com/in/khushg | github.com/khushgx

EDUCATION

University of Pennsylvania   Philadelphia, PA
Jerome Fisher Management and Technology Program (M&T)   *August 2022 - May 2026*
- Bachelors of Science in **Computer Science** and **Finance and Statistics**
- **Relevant Coursework**: Machine Learning, Computer Organization and Systems, Big Data Analytics, Data Structures and Algorithms, Discrete Math, Multivariable Calculus, Probability and Statistics, Linear Algebra

EXPERIENCE

Software Engineering Intern   August 2023 – Present
Cypher Accelerator @ UPenn   *Philadelphia, PA*
- Researching with Federal Reserve and Professor Chris Callison-Burch on Cloud NLP cluster, using Python to develop a machine learning model predicting bank runs based on natural language analysis from Reddit data.
- Managed fintech accelerator, collaborating with Mark Cuban, Bessemer Venture Partners, and A16Z to nurture fintech startups nationwide participants achieved a post-valuation of $2B.
- Spearheaded the development of a dedicated social media application for founder networking, employing JavaScript, Node.js, and MongoDB.

Software Engineering Intern   June 2023 – August 2023
Advanced Health Academy (AHA)   *San Francisco, California*
- Led SDK development in React Native, seamlessly integrating API with EHRs (like Elation) using AWS, DynamoDB, and Python. Published to NPM with 1k downloads and enhanced diagnostic efficiency by 25%
- Integrated the MPT-7B Model with ChromaDB, Python, Langchain to launch an embedding-based chatbot trained on user's blood reports; enhanced blood report interpretation and boosted user query response time by 30%

Software Engineer   Sep. 2022 – August 2023
PALX   *Philadelphia, PA*
- Founded and led a dev team of 10+ at an insights firm, using behavioral economics to enhance foot traffic in hospitality services by 5% encoding decisional matrices to create model in Python, and Go using US Census Data.
- Built and deployed an analytics webapp showcasing US county socioeconomic landscapes using MongoDB, React.js, Docker, Kubernetes, and GCP and Python. Has 1000+ traffic and users per month for Ohio counties

PROJECTS

Discus.ai (Founder) | *Python, Next.js, NoSQL (DynamoDB), Docker, Kubernetes, GCP, AWS, Git, ChromaDB*
- Developed a no-code Synthetic Data Generation Platform to Fine-tune LLMs using Next.js and Python. Utilized AWS, and DynamoDB for backend; deployed on GCP via Docker and Kubernetes. Supports generation from OpenAI, MPT-7B, and LLama-7B, and knowledge base/embedding based data generation
- Created open-source python package on pip for synthetic data generation. Features include data refinement, document-based synthesis, and custom data generation. Has 50 stars, 4 forks, and 250+ users on GitHub

RunBuddi | *C++, Firebase, React Native, Circuits, Git*
- Deployed an XR fitness app using React Native, optimized for Android platforms, incorporating real-time data syncing using Firebase. Displays current pace, last run pace, and target pace based on user data
- Engineered wearable speedometer using C++, mini-circuit boards, LED displays, for real-time speed calculations, integrating GPS data with 3D hand movement vectors. Parsed data to Firebase and achieved a 92% speed accuracy

Spotify Page Rank Recommendation System | *Python, Flask, Docker, GCP, React, CSS, Git*
- Utilized HTML, Tailwind CSS, Flask for UI, deployed via Google Cloud, Docker, Nginx; leveraged Genius & Spotify APIs to gather user insights and song analytics
- Modified Google's Page Rank to recommend songs based on edge weights; vectorized audio/lyrics gathered with cosine similarity for precise recommendations

TECHNICAL SKILLS

Languages: Java, Python, C/C++, SQL (Postgres), Javascript, HTML/CSS, R, Go, OCaml, GraphQL
Frameworks: Next.js, Flask, JUnit, Django, Pytorch, AngularJS, React.js, Tailwind CSS
Developer Tools: Git, Docker, Google Cloud Platform, Amazon Web Services, VectorDBs, MongoDB, Tableau, Excel
Libraries: Pytorch, Pandas, Seaborn, matplotlib, NumPy, Tensorflow, LangChain, LLamaIndex, CUDA