Kartikeya Sharma

LinkedIn | GitHub | jadoo@berkeley.edu | +1 (510) 984 8388

EDUCATION

University of California, Berkeley   Sep 2021 – May 2025
Bachelors' Data Science, Domain Emphasis: Applied Mathematics and Modelling
- Awards and Honors: UWC Davis Scholar ($30,000/annum), CDSS Scholarship ($10,000/annum), OpenAI Hackathon Winner
- Selected Coursework: *Efficient Algorithms and Intractable Problems, Introduction to Machine Learning, Deep Learning for Computer Vision, Data Engineering, Data Structures, Financial Engineering, Optimization for Engineering Models, Probability for Data Science, Linear Algebra*
- Societies: President of IPEF at Berkeley

United World College in Mostar, Bosnia & Herzegovina
- International Baccalaureate Diploma Program (Bilingual Diploma awarded)   Sep 2019 – May 2021

EXPERIENCE

BLCK UNICRN, Los Angeles, CA
Data Science Intern   Sept 2024 – Current
- Analyzed the Talent Marketplace platform and AI Tools, to maximize value for enterprises, SMBs, and creators.
- Designed subscription tier segmentation using data from industry research, aligning offerings with user needs.
- Developed dynamic pricing models in Python, balancing cost, ROI, and competitor pricing.
- Presented and refined pricing strategies using PowerPoint and Excel

Seva Foundation, Berkeley, CA
Data Engineer Intern   May 2024 – Aug 2024
- Implemented a PostgreSQL database from the ground up, managing training data spanning five continents and four languages.
- Built and optimized data pipelines using dbt for robust data validation and cleaning
- Integrated pipeline update notifications with Slack, enabling real-time alerts and improving team communication

Database Systems and Analytics Intern   Jan 2024 – April 2024
- Integrated legacy Excel data into the current Snowflake database, ensuring seamless data migration and continuity
- Developed an internal tool using SQL and Streamlit-in-Snowflake for visualizing surgical quality metrics across vision centres
- Curated data dictionaries for dbt, standardizing data descriptions and enforcing validation protocols

Marvell Nanofabrication Laboratory, Berkeley, CA
Systems Engineering Intern   May 2022 – Aug 2022
- Utilized Python and Pandas for data analysis and automation of engineering tasks
- Contribute to the design and execution of electron beam lithography data to evaluate equipment consistency
- Assist in the implementation of automation solutions to streamline workflows and enhance overall process efficiency
- Engage in rigorous testing and validation of equipment upgrades and modifications

PROJECTS

Steam-lit - *Personalized game recommendation system using advanced NLP techniques*   Jan 2024
- Manually extracted data from Steam Web API with appropriate rate limiting and batch processing
- Implemented BERT for generating context-aware word embeddings of game descriptions to capture nuanced meanings
- Applied LDA to extract dominant topics from user reviews, enriching the recommendation engine with thematic insights
- Integrated BERT matrix with LDA distributions to create a robust design matrix, enhancing the recommendation accuracy
- PCA for dimensionality reduction along with caching for efficient model deployment on Streamlit app
- Link to project: https://steam-lit.streamlit.app

Philosofy - *Debate Simulator using Llama 3.1*   *August 2024*
- Integrated Llama 3.1 advanced natural language processing (NLP) models to handle text generation tasks, ensuring high-quality, context-aware outputs that resonate with users seeking intellectual engagement
- Engineered a scalable backend using Flask, enabling seamless communication between the frontend and the LLM
- Optimized the simulation experience by integrating advanced prompt engineering

Scribble AI – *#29 on Product Hunt, OpenAI Hackathon winner*   June 2023
- Optimized prompt engineering techniques to effectively guide the AI in generating contextually relevant creative content
- Fine-tuning of advanced NLP models using state-of-the-art deep learning frameworks, while handling thousands of concurrent users with minimal latency. Achieved the #27 peak ranking on Product Hunt
- Link to project: https://www.scribbleai.com/

ADDITIONAL INFORMATION

Technologies: Python (PyTorch, PySpark, Scikit-Learn, Transformers, Pandas, Seaborn, Flask, Streamlit, LangChain), SQL (PostgreSQL, SQLite, SQL Alchemy, NoSQL), Java, C# (Unity), AWS, Apache Airflow, Apache Kafka, Snowflake, Docker, Git/GitHub, dbt, Linux