Arpit Parihar

Github | LinkedIn | arpitparihar07@gmail.com | Redwood City, CA | +1.773.837.8826

Work Experience

Staff ML Scientist, Snowflake Inc. Feb 2022 - Present

- Developed an advanced AI-based employee assistant chatbot for IT, Security, Travel, Expenses, and Product documentation, currently used by 9000 employees for day-to-day inquiries:
  - Implemented custom hybrid search from scratch, combining:
    - BM25 lexical search algorithm developed entirely in Snowflake
    - Semantic vector search using a custom-trained e5-unsupervised-small model with contrastive learning developed in PyTorch
    - Cross-encoder reranking and Maximum Marginal Relevance (MMR) algorithm developed in Python and Snowpark to enhance result diversity
  - Utilized open-source LLMs through Snowflake Cortex for response generation
  - Created a user-friendly Streamlit frontend with citations, chat history, and feedback mechanisms
  - Employed RAGAS (Retrieval Augmented Generation Assessment) to quantify and optimize retrieval and response quality
  - Developed an image transcription pipeline using a vision transformer (LLaVa-1.5), enhancing searchability and enabling image retrieval in responses
  - Fine-tuned BERT-based models on custom datasets, enabling guardrails to prevent misuse
  - Currently developing:
    - A knowledge graph-based RAG framework and adapting the solution for customer-facing deployment on the Snowflake website
    - An LLM orchestrator to combine several LLM apps into a single conversational agent

- Worked closely with the corporate security team to identify and mitigate malicious domains generated by domain generating algorithms. Utilized advanced techniques such as Word2vec for IP to DNS path vectorization, cosine distance, and clustering to successfully identify anomalies and improve security measures.

Teaching Assistant, University of Chicago Mar 2021 - Present

- Instructed large classes of up to 60 students in Machine Learning, Python for ML, Reinforcement Learning, Linear & Non-Linear Models, and Machine Learning Operations. Conducted personalized sessions with students to provide individualized guidance and support, resulting in improved academic performance and overall learning outcomes.

Data Science Intern, McKinsey & Company Jun 2021 - Sep 2021

- Conducted experiments using MLFlow and Azure Databricks to build powerful ensemble models that predicted churn among farmers and distributors, achieving an accuracy rate of 68%.
- Took ownership of code refactoring efforts, ensuring that all code was modularized, unit tested, formatted, and optimized for performance across a multi-region team spanning four different time zones.

Senior Data Scientist, Boston Consulting Group Apr 2018 - May 2019

- Led the development of a deep learning model to detect breast cancer from mammograms, achieving an accuracy of 96% using Keras and convolutional neural networks in a proof-of-concept project.
- Designed and implemented a Gradient Boosting model for a leading car seat manufacturer, which was used as an enterprise scale real-time predictive maintenance framework in Python and Azure. The model achieved an F1 score of 76% and resulted in yearly savings of $450k per plant.

Data Science Manager, Flipkart Sep 2017 - Mar 2018

- Led a team in identifying the optimal delivery hub locations for a logistics company in the state of Karnataka using geocoding and hierarchical clustering. The project resulted in a 50-minute reduction in delivery times on average, significantly improving customer satisfaction and overall business efficiency.

Data Scientist, Mu Sigma Mar 2015 - Aug 2017

- Developed customer churn models using logistic regression for a major insurance company, achieving a retention rate of 78%. Involved clients in the analysis journey with infographics and presentations, helping them understand the findings and insights.
- Identified suspicious health insurance claims by formulating hypotheses and developing a rule-based fraud engine. Developed a random forest classifier to prioritize fraudulent claims, resulting in total savings of $11M. The project significantly improved the efficiency of claims processing and saved the company valuable resources.

Education

2020 - 2021 Master's in Machine Learning at **University of Chicago** (GPA: 4.0/4.0)
2010 - 2014 Bachelor's in Civil Engineering at **College of Engineering, Pune** (GPA: 7.5/10.0)