Arnav Shah

+1 667 464 1584 ashah108@jhu.edu arnavsshah.github.io linkedin.com/in/arnav-s-shah

Education

Johns Hopkins University Aug. 2022 – May 2024

Master of Science in Engineering in Computer Science, GPA - 3.97/4.0
Courses - Software System Design, Cloud Computing, Databases, Web Security, Information Retrieval, Blockchain, Natural Language Processing, Machine Translation

Veermata Jijabai Technological Institute (VJTI), India Aug. 2018 – May 2022

Bachelor of Technology in Information Technology, GPA - 9.29/10
Courses - Big Data Analytics, Data Mining & Warehousing, Machine Learning, Operating Systems, Computer Networks, Parallel Computing, Cyber Security, Software Engineering, Algorithms

Technical Skills

Languages: Python, Java, SQL, JavaScript, C++, HTML, CSS

Tools and Technologies: Node.js, React, React Native, Spring Boot, Flask, Docker, PyTorch, TensorFlow, Scikit-learn, Pandas, PySpark, Databricks, Elasticsearch, Apache Airflow, NumPy, Matplotlib, NLTK, AWS (EC2, ELB), Azure (App), RESTful APIs, Kaldi ASR, Git, Bitbucket

Databases: MySQL, PostgreSQL, MongoDB, Neo4j, AWS (S3, RDS)

Work Experience

Johns Hopkins Bloomberg School of Public Health Oct. 2023 – Apr. 2024

Full-Stack Software Engineering Intern

- Led the front-end development to implement interactive driver training and testing modules, achieving a 15% reduction in average page render time.
- Orchestrated the design of REST API endpoints, implemented user authentication, and contributed to database modeling, resulting in a 30% increase in throughput.
- Established a CI/CD pipeline using Microsoft Azure to deploy the application, benefiting over 2000 teen drivers in the state of Washington in anticipating road hazards.

Abbott June 2023 – Dec. 2023

Data Science Intern, Global Data Science & Analytics

- Implemented a data analysis pipeline to identify proxies for assessing the efficacy of Deep Brain Stimulation in Parkinson's patients across over 10 billion medical device records, for submission to FDA and EU regulators.
- Developed interpretable and predictive machine learning models, achieving a 98% accuracy in forecasting heart failure stages by leveraging both insurance claims and EHR data.

Center for Language and Speech Processing, Johns Hopkins University Nov. 2022 – Dec. 2023

Research Assistant, Advised by Prof. David Yarowsky

- Published a Python package for universal voice activity detection, demonstrating competitive performance with a 5% reduction in detection error rate compared to existing open-source models.
- Developed an algorithm to augment Automatic Speech Recognition (ASR) with Grapheme-to-Phoneme models, yielding a 3% improvement in phoneme error rate across 130 languages.

Citigroup May 2021 – July 2021

Full-Stack Software Engineering Intern, Equities Technology Division

- Utilized the QuickFIX/Java library to integrate the "Financial Information Exchange" (FIX) protocol with a regression and load testing tool for Citi's in-house order management system–"One Connect".
- Incorporated WebSockets to pass messages between back-end microservices with an average latency of 60ms.
- Upgraded the JUnit test suite for various FIX message scenarios and brought the code coverage up to 92%.

Projects

DocuSync June 2023

Realtime updates to product documentation using generative AI.

- Managed the back-end infrastructure, database design and LLM prompting using the OpenAI API.
- Contributed to the front-end development of a real-time, editable document-difference service with relevance ranking.
- Received 2nd place at the HackerDojo Generative AI Hackathon, with plaudits in the business use-case category.

Student System May 2020

A web-based platform that links students, educators and alumni, facilitating the exchange of educational materials and fostering community formation.

- Engineered the back-end architecture and database framework, crafting over 80 SQL queries to aid system functionality.
- Built a recommendation engine to track similar students and alumni with a MAP@5 of 94%.