Aditya Pai Brahmavar

Raleigh, NC +1(919)559-7601 theadipai@gmail.com linkedin.com/in/adityapai16 github.com/adipai

Experience

ZS Associates **New York, USA**

Business Technology Solutions Associate Consultant Intern - AI/ML **June 2024 – August 2024**

- Streamlined data model and **optimized SQL queries** for data extraction in Snowflake, achieving 30% faster runtimes while preserving **data integrity** for critical enterprise customer and supply chain data
- Enhanced and scaled a drug shortage prediction solution by generalizing the **machine learning pipeline** using Python and Dataiku DSS, targeting a billion-dollar generics market

Brillio **Bangalore, India**

Senior Data Engineer - Software development **April 2022 – July 2023**

- Built **scalable royalty application** for a leading music corporation, fully orchestrated on AWS to maximize data accessibility for over **1 million artists worldwide**
- Spearheaded the development of an **event-driven data platform** using **AWS Lambda** and **Step Functions**, for real-time generation of royalty statements
- Designed and developed an automated **data validation microservice** with the **Great Expectations framework**, while optimizing for 40% reduced cluster memory usage
- Devised an automated mechanism to synchronize downstream **low-latency databases (Aurora)** and **real-time search (OpenSearch)** with highly updated data from the **data warehouse (Redshift)**

Data Engineer - Software development **September 2020 - March 2022**

- Automated the migration of on-premise data storage to **data lake** on **AWS S3** using **Storage Gateway**, reducing operational costs by over 50% for a US-based power company
- Developed **ETL pipelines** for the energy efficiency initiative, boosting customer retention by 20% with targeted analytics that guided customers on reducing electricity consumption
- Implemented **large-scale predictive analytics and machine learning** using PySpark to predict electricity outages, aimed at reducing time to restoration by 50%

Skills

Programming	Python, Java, C++, HTML, CSS, JavaScript, Linux
Data & AI	PostgreSQL, MySQL, Apache Spark, Elasticsearch, Databricks, Dataiku DSS, Snowflake, Scikit-learn, MLlib, Keras, Pytorch, Tensorflow, MLflow
Cloud & Web	Microsoft Azure (2x Certified), AWS, REST APIs, FastAPI, Django
Tools & DevOps	Docker, Kubernetes, Terraform, Git, Postman, Spring Boot
Methodologies	Object Oriented Programming, Agile SDLC, Test Driven Development, Unit Testing

Education

Master of Computer Science **August 2023 – May 2025**

North Carolina State University, Raleigh, NC **GPA: 4.00/4.00**

Coursework : Algorithms, Software Engineering, Graph Data Mining, Computer Vision, Resource-dependent Deep Learning

Bachelor of Engineering in Information Science **August 2016 – May 2020**

Nitte Meenakshi Institute of Technology, Bangalore, India **GPA: 8.73/10.00**

Coursework : Data Structures, Operating Systems, Parallel Computing, Distributed Systems

Projects

Web application for educational management

- Developed **concurrent APIs** with Python and **FastAPI**, integrated with a TypeScript frontend, to enable distinct capabilities for each user role on the platform
- Implemented **database design** and role-based **CRUD operations** for tasks like managing courses, tracking student progress and facilitating course material access

Web application for cryptocurrency price forecasting | Publication

- Built **Flask**-based web application to generate price forecasts for cryptocurrencies over user-defined time frames
- Trained **LSTM models** with MSE of 0.000001 units for real-time inference

Graph Data Mining for Distributed Application Security | Publication

- Developed framework for graph representation of microservice architecture to model access patterns and user behavior
- Applied supervised ML to detect security threats through real-time **graph analysis**

Web application for Pneumonia detection

- Created Flask-based web application for Pneumonia detection from chest X-rays with 95% accuracy
- Trained **CNN model** on 1000+ labeled images sourced from hospitals, for backend processing and inference