

Adith Prabukumar

📞 7169944271 | ✉ adithprabu27@gmail.com | [in adithprabukumar](https://www.linkedin.com/in/adithprabukumar) | [github adith-prabukumar.github.io](https://github.com/adith-prabukumar)

WORK EXPERIENCE

Frontier Science And Technology Research Foundation

March 2023 – Present

Data Engineer

Remote, USA

- Designed **ETL Framework** using **dbt** and **Apache Spark**, to transfer data from hospitals to AWS cloud servers, leading in 60% in employee productivity.
- Collaborated with Senior Enterprise Architects, Solution Architects, Product Managers, to understand the workflow and ensure standards and rationalization plans to improve visibility and insights of strategic outcomes to drive product **impact by around 70%**.
- Worked** on real-time data transfer pipeline utilizing **Kafka**, transferring over **1M patient data** from local server to **AWS Redshift** data warehouse, ensuring ease of access to data and increasing productivity by 80%.
- Played a pivotal role in migration of complex data sets to AWS deploying **AWS DMS**, establishing easy access to data for more than 3000 employees across the different parts of the country.
- Developed a customized **NLP chatbot** for employee inquiries utilizing **Large Language Models (LLMs)** and integrated **Snowflake** Security tool for data access control based on clearance levels, boosting productivity by 60%.
- Worked on unit tests using **JUnit** and performed quality checks to achieve software functionality, glitch elimination and update maintenance of access control, leading to **70% increase in performance**.
- Employed **Apache Airflow** to execute DAGs for data collection, reducing data processing time by 67%.
- Spearheaded a design and implementation of robust **CI/CD pipelines** using **GitHub Actions** and other version control methodologies across all projects.

AINQA GROUP

August 2020 – July 2021

Data Analyst

Tamil Nadu, India

- Built a robust website using **Python and Flask** capable of parsing CSV files and pushing to the company's database.
- Leveraged **OCR technology** using **OpenCV** to extract structured data from clinical files from pharmacies leading to the creation of a comprehensive data set, increasing the productivity of the team by 40%.
- Developed and maintained interactive **Tableau** dashboards that visualized critical business metrics for executive decision-making, which was presented to Senior Product Managers and Company Vice President.
- Employed diverse **KPI trackers** and vendor graphs to refine the codebase, resulting in optimized performance while creating **user-friendly documentation** for analysts.

EDUCATION

University at Buffalo, The State University of New York

August 2021 – February 2023

Master of Science in Data Science

Buffalo, NY

Cumulative GPA: 3.8/ 4.0 - Magna Cum Laude

Anna University

August 2017 – February 2021

Bachelor of Mechanical Engineering

Tamil Nadu, India

Cumulative GPA: 3.6/ 4.0

SKILLS

Certifications : **DataBricks** Certified Data Engineer Associate

Languages : Python, R, Go, C, Java, SQL (PostgreSQL, MySQL), NoSQL, Git

Big Data : Apache Hive, Kafka, Hadoop, Apache Spark, Apache Airflow, Hadoop.

ACADEMIC PROJECTS

News Summarizer: Developed a ML model that summarizes news articles using Python, Docker and AWS. [Click Here](#)

Audiolytics: Developed an end-to-end ML model that implements speech-to-text on audio files using Python, Docker and Github Actions. [Click Here](#)

2020 Data Science Job Analysis: Applied NLP-based transformation on data science job data scrapped from Glassdoor using Python and Tableau. [Click Here](#)