

# Adith Prabukumar

☎ 7169944271 ✉ adithprabu27@gmail.com in LinkedIn

## Professional Experience

**Data Engineer, Frontier Science and Technology**

Aug 2022 | Buffalo, NY

- Designed **ETL pipeline** using **dbt** and **Apache Spark**, to transfer data from hospitals to AWS cloud servers, leading in 60% in employee productivity
- Developed a dashboard that analyzes real-time query data and database performance using Tableau, improving data visibility by 60%
- Implemented **optimized PostgreSQL queries** on legacy projects, decreasing query processing time by 40% and improving cross-team functionality.
- Automated **data retrieval** process using **Google Sheets API** modules in Node.js, which reduced the access time by 40%
- 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%
- Worked on unit tests using **Unittest** 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.

**Data Engineer, AINQA Group**

Aug 2020 – Jul 2021 | Mumbai, India

- Built a 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

## Skills

**Languages:** (Python, Go, Javascript, HTML, Bash, Git)

**Python Packages** (Pandas, Numpy, Django, Flask, FastAPI)

**Database** (PostgreSQL, MySQL, MongoDB, Oracle SQL, Firebase)

**Tools** (Github Actions, Grafana, Tableau, Postman, Docker, Kubernetes, AWS, GraphQL)

**Frameworks** (React.js, Flask, Django, Node.js, Pyspark)

## Education

**Masters of Science in Data Science, University at Buffalo**

Aug 2021 – Feb 2023 | Buffalo, NY

**Cumulative GPA:** 3.8/ 4.0 - *Magna Cum Laude*

**Bachelor of Engineering, Anna University**

Aug 2017 – Mar 2021 | Tamil Nadu, India

**Cumulative GPA:** 3.7/ 4.0

## Projects

**Khoj, Active Open-Source Contributor** ✉

Aug 2023 – present

- Leveraged knowledge of **Object Oriented Programming** concepts to write clean code and improve functionalities.
- **Resolved 4 open issues** to completion. Collaborated with various other contributors to implement the Github indexing API feature.
- Played a key role in the **documentation** of the new software changes and feature updates.