

Adith Prabukumar

📞 7169944271 | ✉️ adithprabu27@gmail.com | [in adithprabukumar](https://www.linkedin.com/in/adithprabukumar) | [Adith2710](https://github.com/Adith2710)

WORK EXPERIENCE

Frontier Science And Technology Research Foundation

March 2023 – Present

Data Engineer

Remote, USA

- Developed a **bot** that annotates case report forms extracted from PDF files leading to **63.2% increase** in streamlining data annotation.
- Executed the processing of clinical data following FDA standards, utilizing **SQL DDL and DML commands**.
- Managed workflows to transfer data to subsequent stages within the data management pipeline, ensuring data integrity and compliance.
- Played a pivotal role in the successful migration of complex data sets to AWS using **AWS Data Pipeline**, resulting in a 33% reduction in computing time, optimizing data processing efficiency.
- Collaborated closely with study build teams from Harvard University to facilitate code checks across various domains.
- Employed **Apache Airflow** to implement DAGs for data collection, executing automated data processing scripts for streamlined data validation.
- Spearheaded the 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 Engineer

Tamil Nadu, India

- Engineered an **ETL pipeline** that migrated data from prototype to production stages reducing data entry efforts by **80%**, optimizing allocation.
- Built a robust website capable of parsing Excel and CSV files, ensuring compliance with the company's SQL database standards.
- Implemented **Python scripts** to automatically detect and handle outliers, resulting in a remarkable 45% increase in productivity.
- Leveraged **OCR technology** to extract structured data from clinical files obtained from pharmacies. This implementation led to the creation of a comprehensive and organized clinical data set, facilitating data-driven decision-making.
- Innovated a **Tableau**-based dashboard tailored for vendors, offering comprehensive insights into product performance and trends.
- Incorporated various KPI trackers and vendor-specific graphs, empowering vendors with actionable business insights.
- Designed and implemented a model architecture, integrating transfer learning techniques to achieve an exceptional accuracy rate of 85%.
- Developed the solution using Python tools including **PyTorch, NumPy, and OpenCV2**, highlighting expertise in advanced machine learning methodologies and proficient utilization of essential libraries for robust model creation.
- Methodically investigated previous projects that utilized the company database, strategically enhancing **SQL queries to boost performance**.
- Implemented optimization techniques resulting in an average reduction of query execution time by 30%, leading to more efficient data retrieval and improved overall system responsiveness.

Specrom Analytics

December 2019 – June 2020

Data Science Intern

Maharashtra, India

- Engineered a robust machine learning model by harnessing the **BERT transformer** architecture to address essential **NLP tasks**.
- Employed **Keras** to meticulously fine-tune gradients, significantly enhancing model performance through strategic training on a meticulously curated dataset of news articles.
- Spearheaded the establishment of a CI/CD pipeline through GitHub Actions, resulting in swift and dependable deployment of the project and its subsequent updates. This strategic implementation streamlined development workflows, enhancing project efficiency.
- Led the design and implementation of a comprehensive end-to-end pipeline for Amazon product review scraping, optimizing data collection efficiency and accuracy.
- Developed a user-friendly Tableau dashboard tailored for vendors, providing in-depth insights into product performance, including key performance indicators (KPIs) and vendor-specific graphical representations.
- Conducted in-depth analysis on Walmart product reviews within the watch category, extracting valuable insights into purchase patterns and user feedback behaviors to inform strategic decision-making.

EDUCATION

University at Buffalo, The State University of New York

August 2021 – February 2023

Master of Science in Data Science

Buffalo, NY

Cumulative GPA: 3.76/ 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 Data Engineer Associate, AWS Certified Cloud Practitioner

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

Big Data : Apache Hive, Kafka, Hadoop, Spark