Raashid Salih

rashidsaleh111@gmail.com +91 62-826-22077 raashidsalih.github.io

EDUCATION

2023 – present
Atlanta, Georgia, USA

Georgia Institute of Technology (GaTech) [Currently on Academic Hiatus]
Master of Science in Analytics (MSc)
- CGPA - 4.00/4.00

2018 – 2022
Kuala Lumpur, Malaysia

International Islamic University Malaysia (IIUM)
Bachelor of Computer Science (Honours)
- ***Specialization in Computational Intelligence and Data Science***
- CGPA - 3.99/4.00

PROFESSIONAL EXPERIENCE

Dec 2023 – present
Bangalore, India

FreightFox (Logistics based AI Startup)
Data Scientist
[*Python*, *AWS*, *Docker*, *FastAPI*, *Postman*, *Streamlit*, *GenAI*, *Qdrant*, *Redash*, *PuLP*, *Jira*, *Confluence*]
- Shaped company's Data Science strategy by curating features and aligning development with product roadmap for maximum impact.
- Engineered industry-first toll-to-toll latency index with optimized SQL and Python to handle large datasets to help clients with logistics planning and notify them of network latency changes.
- Created comprehensive anomaly detection suite which reduces client cognitive burden by highlighting critical actionable shipments.
- Developed RAG based LLM solution (Text2SQL) using to facilitate self serve analytics for users via chat interface.
- Markedly enhanced data quality by implementing novel solution to industry-wide problem of non-standard location names enabling inter-manufacturer analytics.
- Leveraged said solution to automate client logistics network creation, significantly reducing onboarding time to improve client acquisition.
- Productionized a demand forecasting pipeline to aid in inventory management.
- Innovated a vehicle categorization method utilizing application-specific metrics with one-dimensional clustering, solving industry-wide issue of non-standard vehicle names.
- Deployed a web application to visualize commodity trends and sentiments from the news to help clients understand commodity risk and engage in proactive supply chain management.
- Built an ensemble model to predict freight rates based on shipment features for clients to benchmark and optimize performance.
- Developed an optimization model to minimize costs by determining the optimal fleet mix and payloads given a total payload.

Jan 2022 – Sep 2022
Kuala Lumpur, Malaysia

Schlumberger (Energy MNC)
Data & Analytics Intern
[*Python*, *Dash*, *Dataiku*, *Azure DevOps*, *SSMS SQL*, *Alteryx*, *PowerBI*, *Dax*]
- Enhanced predictive model reliability by applying Explainable ML techniques leading to identification of model strengths and weaknesses.
- Developed a cost prediction WebApp enabling project cost forecasting based on inflation rates.
- Implemented pipelines for sprint planning and department KPO dashboards using ensuring comprehensive documentation and adherence to best practices, improving data visibility.

PROJECTS

End-to-End Churn Analytics Platform
[GCP, Docker, Airflow, dbt, PostgreSQL, Terraform, FastAPI, MLFlow, Grafana, Metabase, FLAML AutoML, Ruff]
Developed a cloud-hosted containerized data platform to simulate and analyze customer behavior utilizing microservices architecture. Integrated ML models for synthetic data generation and churn prediction served via RESTful APIs. Ensured data quality with automated tests and transformations, and designed an interactive dashboard for key metrics. MLOps best practices like experiment tracking and monitoring is facilitated by MLFlow and Grafana. Implemented CI/CD for seamless updates and provided extensive documentation for reuse.

Agmarknet Scraper
[Python, ArgParse, PostgresSQL, Selenium, dotenv, venv]
A CLI tool designed to gather data from Agmarknet . The current implementation employs a direct URL substitution approach to minimize interactions with the website for stability, and moves the relevant data collected to a Postgres database (as default behavior). The Python project uses the Model-View-Controller (MVC) architectural pattern for better separation of concerns and isolates database connection details in an .env file for greater security.

JenksPy: Fisher-Jenks algorithm (Python)
Open source contribution made to the Fisher-Jenks algorithm implementation to help determine ideal class/cluster size using elbow charts based on the goodness of variance fit (GVF) metric.

MISCELLANEOUS

- **IELTS Academic: Band 9 | CEFR - C2 (highest level of English proficiency)**