

Data Engineer Intern at AIDF-CRI

Background

The Asian Institute of Digital Finance (AIDF) is a university-level institute in NUS, jointly founded by the Monetary Authority of Singapore (MAS), the National Research Foundation (NRF) and NUS. AIDF aspires to be a thought leader, a Fintech knowledge hub, and an experimental site for developing digital financial technologies as well as for nurturing current and future Fintech researchers and practitioners in Asia. The Credit Research Initiative (CRI) is a non-profit undertaking under the AIDF. Pioneering the "public good" credit risk measures, the initiative is committed to advancing big data analytics and providing directly useful credit intelligence to academic and professional communities.

Moreover, AIDF-CRI is dedicated to staying updated with the latest trends and technologies, especially for AI and LLMs. We are currently in the process of productionizing an LLM-driven application system to enhance and diversify our credit solutions. This includes redesigning, testing, and improving the backend framework.

Responsibilities

AIDF-CRI is actively seeking candidates with a deep passion for backend development. The selected candidate will play a key role in providing cutting-edge modern credit analytic platforms by upgrading our data pipeline to meet professional, market-ready standards, focusing on efficiency, functionality, and user experience.

Particularly, the responsibilities will include:

- Data Pipeline Development
 - Design and build ETL (Extract, Transform, Load) or ELT (Extract, Load, Transform) data pipelines using BigQuery and other Google Cloud Platform (GCP) services.
 - Develop and optimize data models and schemas to support analytics, reporting, and machine learning requirements.
 - Automate workflows using tools like Cloud Composer or Cloud Functions to ensure reliable data pipelines.
- Team Collaboration
 - Collaborate with credit analysts, data engineers, and other stakeholders to ensure seamless migration of data treatment and modeling logic from legacy systems.
- Optimization and Testing
 - Establish and enforce data quality standards and governance frameworks to ensure data integrity and security.
- Documentation
 - Create comprehensive documentation for data pipelines, workflows, and processes to promote knowledge sharing and ensure smooth onboarding for new team members.

Minimum Requirements

- Preferably major in the areas of computer science/engineering or similar discipline.
- Familiar with at least one programming language (Python, Java, Golang).
- Experience working with database servers, such as MySQL and MongoDB.
- Familiarity with version control systems like Git.
- Strong analytical and problem-solving skills, with the ability to troubleshoot and resolve issues effectively.
- Excellent communication and documentation skills, both written and verbal, to collaborate effectively with team members and stakeholders.

Bonus Skills

- Experience with Docker, Kubernetes and Redis.
- Experience working with Google Cloud Platform, Airflow and Google BigQuery..
- Ability to think creatively and propose innovative solutions to complex problems.