Jiajun Liu

Miami, FL | 786-683-9715| jiajunliu111@gmail.com | www.linkedin.com/in/liujiajun/

PROFESSIONAL SUMMARY

Data professional with 5 years of experience across analytics, data engineering, and ML. Lead initiatives in Airflow pipelines, GBM models, and executive reporting. Skilled in SQL, Python, Power BI, and Salesforce CRM Analytics..

TECHNICAL SKILL:

Languages & Analytics: Python (Pandas, NumPy, scikit-learn, Spark), SQL, DAX, M Code
Machine Learning: Linear models, Clustering, GMM, PCA, ICA, RP , OpenAI embeddings models**.**
Data Platforms: Microsoft Fabric, Azure OneLake, AWS Starburst, Salesforce, Oracle, Postgres SQL
Tools: Power BI, Tableau CRM, Excel, GitHub, JIRA, VS Code, Ubuntu, Virtual Machines, Airflow

WORK EXPERIENCE

Macquarie Capital
Executive – Data Analytics | Senior Associate – Data Analyst | Associate – Data Operations Analyst
May 2021-Present | Jacksonville, FL
Promoted through three roles for delivering end-to-end data solutions that enhanced operational efficiency, strengthened data integrity, and empowered business stakeholders with scalable analytics. My responsibilities evolved from operational data management to advanced analytics, and now to leading data engineering and data science initiative.

- **Data Strategy & Stakeholder Engagement**.
  Partnered closely with senior stakeholders to gather requirements, analyze business processes, and translate them into impactful reporting and data solutions. Led development of dashboards focused on Director KPIs, financial analysis, client activity, and CRM engagement.
- **Data Analytics & Reporting**
  Designed and automated high-impact dashboards using CRM Analytics and Power BI, enabling self-service insights and reducing over 100+ hours of manual reporting effort. Implemented Power BI REST APIs to refresh dashboards only when upstream data updates occurred, optimizing resource usage.
- **SQL Engineering & Data Integration**
  Connected and transformed internal and external data sources into **AWS Starburst** and **Impala Cloud** environments. Wrote and optimized complex SQL scripts to support reporting across disparate systems, enabling unified and reliable analytics across teams.
- **Python & Automation**
  Utilized **Pandas, NumPy, and other Python libraries** for data preprocessing, feature engineering, and validation to support machine learning workflows and advanced reporting. Created Python-based ETL jobs and batch processes to enhance scalability.
- **Data Governance & Security**.
  Defined row-level security predicates in Salesforce CRM Analytics to ensure appropriate data access. Conducted regular audits and implemented data validation rules to uphold data quality and compliance standards.
- **AI & Innovation Initiatives**
  Led an award-winning **AI Hackathon**, leveraging Microsoft Azure OpenAI APIs to embed and index internal documents, resulting in a significantly improved internal knowledge search system and reduced lookup time for documentation.
- **Data Engineering & Machine Learning (Executive Role)**
  Transitioned into a data engineering and data science role. **Architected automated data pipelines** using **Apache Airflow** to ingest, transform, and deliver real-time data to internal platforms. Developed and deployed **gradient boosting machine (GBM)** models to analyze market trends and investor behavior, directly supporting capital raising strategy and client segmentation. Established **feature stores**, model monitoring pipelines, and automated alerts for upstream schema or data quality issues, ensuring robust ML deployment lifecycle.

EDUCATION

Georgia Institute of Technology, College of Computing   **Aug 2023 – Dec 2025 (Expected)**
Master of Science in Computer Science with focus on Machine Learning
University of Florida, Warrington College of Business   **May 2021**
Bachelor of Arts in Business Administration   Gainesville, FL