

Dear Hiring Manager at Data Frameworks,

As a skilled graduate student in Big Data, Cloud, and Software Development, I am excited to apply for the Data Engineer I position in the Enterprise Data Platforms, Risk and Finance Technology team. With a strong background in Python, SQL, and AWS, I am confident that I can deliver high-performing data and analytical solutions that transform, integrate, and make data available for business.

Here are some key skills and qualities that I possess, as demonstrated by my actual experiences:

- Building technical solutions for Data Frameworks:
 - Developed a data pipeline leveraging Apache Spark to process cricket match data, handling large-scale data ingestion from Amazon S3, transforming and validating schemas, and optimizing performance for fast querying and analytics (Sports Data Processing Pipeline).
 - Engineered scalable cloud-based image recognition services using AWS, implementing an advanced auto-scaling algorithm (Elastic Image Recognition Service on AWS).
- Identifying opportunities to expand frameworks and create new frameworks:
 - Led data cleaning and preprocessing initiatives for the house price prediction project, utilizing Python to handle missing values, outlier detection, and feature engineering, resulting in a model with 20% improved pricing accuracy in a tight timeline (Tequed Labs).
 - Conducted comprehensive exploratory data analysis on large-scale logistics datasets, uncovering critical KPIs that were strategically integrated into Tableau dashboards, driving enhanced decision-making (Cube Logistics).
- Building robust data pipelines to collect, process, and compute different metrics:
 - Extracted, analyzed, and visualized reports from Workday and PeopleSoft in Tableau to enhance data-driven decision-making (Arizona State University).
 - Developed a real-time data processing pipeline using Apache Kafka for streaming stock market data, AWS S3 for storage, and AWS Glue to transform raw stock market data into a queryable format and automate the loading process (Real-Time Stock Market Data Pipeline).
- Designing application architecture and technical design:
 - Engineered VBA-embedded Python scripts that automated the manual process of formatting accounts data, achieving a 97.22% reduction in processing time from 12 hours to 20 minutes (Arizona State University).
 - Orchestrated and optimized data pipelines using scikit-learn, incorporating GridSearchCV for hyperparameter tuning, achieving price prediction accuracy of 91% (Tequed Labs).
- Collaborating with cross-functional team members and deploying high-performing code:
 - Coordinated with business users for User Acceptance Testing and with Operations team for code deployment to upper environments (Arizona State University).
 - Deployed high-performing, quality/bug-free code in production, executing unit tests of data populated in target data containers, validating expected results, and ensuring quality and accuracy (Arizona State University).

I am excited about the opportunity to bring my skills and experience to the Enterprise Data Platforms, Risk and Finance Technology team and contribute to building high-performing data and analytical solutions.

Sincerely,
Harsh Victor Challa