

Albert Joshua Capistrano

Westfield, NJ

908-209-0019 • capistranojosh@yahoo.com

SKILLS

- Advanced proficiency in:
 - Amazon Web Services including Lambda, S3, SNS, Redshift, CloudWatch, EMR
 - Python Programming along with Pandas, Boto3, PySpark
- Intermediate proficiency in:
 - SQL queries using Oracle SQL Developer
- Experienced with Git and Github
- Involved working in Agile Environment

EXPERIENCE

Bloomberg L.P., New York, NY

May 2023-Present

Data Engineer

- Designed and implemented Spark-based ETL pipelines to efficiently deliver data to clients on Bloomberg Terminal, ensuring timely and accurate data delivery
- Developed and maintained microservices similar to AWS Lambda integral to ETL pipelines, facilitating preprocessing tasks and enhancing data quality
- Actively participated as a representative of the technology team in the department's knowledge council, contributing to the creation of comprehensive documentation on the tech stack used in projects
- Collaborated closely with team representatives to understand and address business needs, ensuring that technical solutions aligned with client requirements and expectations
- Played a key role in bridging the gap between technology and business by translating business requirements into technical specifications and delivering solutions that meet client needs effectively
- Designed and implemented post-processing microservices tailored to align with specific business requirements, ensuring data quality and integrity before delivery to clients
- Collaborated with stakeholders to understand business objectives and translated them into technical requirements for post-processing microservices, ensuring alignment with client needs

Fannie Mae, Reston, VA

January 2020-May 2023

Software Engineer

- Designed and developed scalable AWS infrastructure to store and process large data sets using Pyspark, S3, EMR, and Redshift
- Implemented ETL pipelines using Amazon Web Services and Python to extract data from various sources, perform data transformations, and load data into the data warehouse
- Utilized AWS EC2 instances to create and manage EMR clusters for data ingestion, ensuring optimal performance and resource utilization
- Implemented Spark jobs on AWS EMR to process large datasets, improving data ingestion efficiency and reducing processing time
- Developed custom scripts to automate data ingestion, data quality checks, and data reconciliation processes
- Utilized AWS Glue for scheduling and running ETL jobs, reducing manual intervention and increasing reliability
- Collaborated with cross-functional teams to gather requirements, design solutions, and implement end-to-end data pipelines
- Monitored performance and optimized AWS resources usage to reduce costs and increase efficiency
- Automated deployment of Lambda functions using AWS CloudFormation
- Designed, developed, and deployed custom software applications using languages such as Python, and Java

EDUCATION

Rutgers University, School of Engineering, New Brunswick, NJ

December 2019

Bachelor of Science in Electrical and Computer Engineering

GPA: 3.45/4.00

- Concentration in Computer Engineering
- Honors: Magna Cum Laude
- Dean's List: Fall 2017, Fall 2018, Fall 2019