Prajeeth Emanuel

emanuel1025@gmail.com | (631) 202-9716 | San Francisco, CA

WORK EXPERIENCE

Amazon.com Services LLC

Software Development Engineer II   May 31st 2022 - Present

- **Architected and implemented** highly reliable backend services for Amazon Search, managing 1.5 billion daily queries with a 99.9% response time of under 150ms, utilizing AWS ECS, SQS, S3, DynamoDB, and Java.
- **Reduced operational costs** by 25% by optimizing data storage solutions and refining service architecture to minimize computational overhead.
- **Led a team of six engineers**, coordinating the integration and deployment of critical projects across Amazon's Retail and Search organizations, enhancing system performance and reliability.
- **Designed and implemented** batch data ingestion pipelines for recommendation quality control, filtering, enrichment, and workflow automation, handling 10,000+ updates per second with near real-time processing using Apache Spark, and Apache Airflow.

Apple Inc.

Software Engineer   Jan. 2019 – September 2020

- **Scaled batch ETL jobs** from prototype to 50+ TB/day, enabling real-time information retrieval of webpages, by developing terabyte-scale distributed data processing systems using Spark, Kafka, and Java.
- **Redesigned data architecture** for optimal storage and retrieval of terabytes of data (real-time and batch streaming), cutting storage costs by 4x with Apache Cassandra, Solr, Redis, and Parquet.
- **Led a team of four engineers** and cross-collaborated with multiple teams & stakeholders over a period of twelve months building the legally compliant platform.

Qalaxia Inc.

Software Engineer, Data   Jul 2016 – Dec 2018

- **Designed and implemented** large stateful batch and real-time data pipelines to mine terabyte scale datasets using Spark, Kafka, Java and stored processed data in MongoDB.
- **Built high-performance ETL pipelines** and data warehousing systems to enable data science workflows and improved ingestion of 10 million structured records and processing time by 30%.

PUBLICATIONS & GRANTS

- **ICDCS** (2022) - Near-Data Processing For Analytics Frameworks
- **ACM SIGSPATIAL** (2021) - GPU-Based Real-Time Contact Tracing at Scale
- **NSF SBIR** (2018) [#1843326](#) - Skill-aware query engine for K12 Classrooms

EDUCATION

Stony Brook University   New York, NY
Master of Science in Computer Science; 3.92 GPA   *2021 - 2022*

IIIT Hyderabad   Hyderabad, IN
Bachelor of Technology in Computer Science   *2012 - 2016*

TECHNICAL SKILLS

- **Languages:** Java, Scala, Python, JavaScript, C++
- **Distributed Systems Frameworks:** Flink, Spark, SQS, Kafka
- **Monitoring Tools:** Grafana, Splunk
- **Datastores:** MySQL, MongoDB, Solr, DynamoDB, Elasticsearch, Cassandra, Redis
- **Infrastructure & Workflow Scheduling:** AWS, GCP, Jenkins, Airflow