Lakshmi Pooja Devarapu

Bayonne, NJ (980) 382-4746 ‡ GitHub LinkedIn Portfolio devarapupooja999@gmail.com

Summary

A master's student in Computer Science with a focus on data engineering, offering 1.5+ years of practical experience in developing and managing data pipelines and cloud-based data warehousing solutions. Skilled in leveraging AWS, Python, Kafka, and Spark to build scalable data systems.

Education

University of North Carolina at Charlotte, NC, MS in Computer Science in Data science — Aug 2022 - Dec 2023

Skills

Java, Python, C, Scala, MySQL, MongoDB, PostgreSQL, NoSQL, Snowflake, Amazon Redshift, Apache spark, Kafka, Datalake, Hadoop, Eclipse, Jupiter Notebook, Visual Studio, AWS, Azure, Docker, Kubernetes, Azure Devops, Talend, Apache Airflow, Aws Glue, PowerBI, Tableau, Rest APIs, Web Services, Unix, Linux, MacOS, Windows, Git, GitHub, Bitbucket, Agile, Scrum

Experience

Data Engineer ,
United Airlines, TX — Feb 2024– Present

- Designed and implemented data ingestion solutions using **Hadoop**, **MapReduce**, **HBase**, **Hive**, **Spark**, **AWS S3**, **AWS Lambda**, **DynamoDB**, **EMR**, **Redshift**, and server management **APIs**.
- Engineered scalable multi-tiered web services using **Java** and **MongoDB**, achieving a data processing throughput increase of 30% while managing deployment with **Docker** and **Kubernetes** for improved operational efficiency.
- Orchestrated the development of real-time Tableau dashboards, enabling 15+ team members to access key metrics and enhance decision-making across departments with actionable insights using **Python, SQL, Scala, Snowflake, and Spar**k.
- Designed and optimized 15+ **data models** and **warehousing** solutions, developed and maintained 30**+ ETL pipelines** in **SQL**, and implemented secure data access platforms, resulting in a 40% improvement in reporting efficiency
- Demonstrated proficiency in relational and non-relational databases, object-oriented scripting **(Python, SQL, TypeScript, Java), and REST APIs** enhancing data accessibility for 15+ team members.

Data Engineer ,
Varky Technology Solutions, Hyderabad, IN — June 2021 – Jun 2022

- Developed and optimized 25+ data pipelines using **AWS, Python**, **SQL**, and **Apache Spark**. Skilled in **Snowflake** in cloud-based warehousing and **Kafka** for real-time data streaming, improving data processing efficiency by 35%.
- Orchestrated 40+ **data workflows** using **Azure Data Factory** (ADF) and **Airflow**, automating integration and transformation processes to enhance data pipeline efficiency by 30%.
- Utilized AWS technologies (**Redshift, S3, Glue, EMR, Kinesis, Firehose, Lambda, IAM**) for scalable data architecture, and Engineered solutions using non-relational databases (object storage, document/key-value stores, graph databases) resulting in **a 35% improvement in query performance**
- Enhanced application and query performance by 25% using profiling tools and **SQL**, while implementing best practices in **data integrity**, analysis, and **documentation** to ensure robust and reliable **data management**.

Projects

COVID 19 Big Data Analysis

- Outlined and managed data engineering solutions using **Azure Data Factory**, **Data Lake Gen2**, **Blob Storage**, and **Databricks**, including pipeline integration and complex control flows.
- Applied and monitored **ADF pipelines** with **Azure Monitor**, configured **CI/CD** processes using **Azure DevOps**, and managed **scheduling** and **debugging** for data transformations, resulting in a 20% reduction in deployment time.

Malaria identification using a neural network

- Originated and executed a **CNN** model analyzing 10,000+ blood microscopy images, reducing detection time by 40% for lab technicians.
- Leveraged **deep learning** for automated feature **extraction**, boosting diagnostic accuracy by 30% and addressing complex challenges.

Certifications

- AWS certified cloud practitioner
- Azure Databricks attested Apache associate spark developer
- Azure Databricks licensed lake house fundamentals
- Udemy accredited in Snowflake master class
- Machine learning in Python (IBM issued
- PET and BEC certified in English (Cambridge issued).