

Niranjan Kumar Gurram

Email: niranjangurramk@gmail.com |

Contact: +1 (682)203-6753

| LinkedIn ID: linkedin.com/in/niranjangk

SUMMARY

Software Engineer with 7+ years of professional experience building scalable cloud-native applications, data engineering pipelines, and backend systems across telecom, analytics, and enterprise domains. Strong expertise in designing and automating end-to-end ETL and data processing workflows using Python, SQL, Apache Spark (JVM-based), and Apache Airflow, supporting high-volume and real-time data workloads. Experienced in full-stack and backend development using React, JavaScript, Node.js, Express, Flask, and Laravel to deliver reliable RESTful APIs and services. Hands-on experience with Java-based big data frameworks within the Hadoop and Spark ecosystem for distributed data processing. Proficient in managing relational and NoSQL databases including MySQL, PostgreSQL, MongoDB, and Azure Cosmos DB, with strong understanding of OLTP and OLAP architecture, data modeling, and governance. Skilled in cloud platforms (Azure, AWS), Infrastructure as Code (Terraform), and CI/CD automation using Jenkins and GitHub Actions to deliver secure, production-ready solutions. Experienced working in Agile environments, focused on system reliability, performance, and maintainability. Holds a master's degree in computer science and information systems from The University of Texas at Arlington.

SKILLS

- Programming Languages: Python, SQL, Java, C, C++, JavaScript.
- Web Technologies: HTML5, CSS3, React, Node.js, Express, NPM, Laravel, Webpack, Bootstrap.
- Database: MySQL, PostgreSQL, MongoDB, Azure Cosmos DB, OLAP, OLTP, DDL, DML.
- Data Engineering & Big Data Technologies: Apache Spark, Apache Airflow, Hadoop, MapReduce, ETL Tools.
- Analytical Tools: Tableau, Power BI, Spotfire, Excel.
- Machine Learning & AI: ML, Deep Learning, TensorFlow, PyTorch, Scikit-learn, CNN, RNN, LSTMs, Transformers
- Python Framework: Flask.
- Familiar Technologies: ServiceNow, AWS, Terraform, Jenkins, Azure Monitor, Application Insights, SHA 256 Algorithm, Apache Cassandra, Google BigQuery, Snowflake, Neo4j, Lucidchart, Draw.io, Django, UML, ER Diagrams, XML, JSP.

WORK EXPERIENCE

Sr. Software Engineer, SAAKHIL Technologies Pvt Ltd. - India.

July 2019 - Aug 2023

Product: Cloud-Native Full-Stack Application & DevOps Automation Platform

Tools: React.js, JavaScript, Node.js, Express, MySQL, PostgreSQL, Azure Cosmos DB, Microsoft Azure, Terraform, Jenkins, GitHub

- Led design and development of cloud-native full-stack applications using React.js, Node.js, and Express, supporting high-traffic enterprise workloads.
- Architected and implemented RESTful APIs with performance tuning, caching strategies, and optimized data access patterns to ensure stable response times under peak load.
- Designed and optimized relational database schemas and queries in MySQL and PostgreSQL, supporting transactional workflows and downstream analytics.
- Automated Azure infrastructure provisioning using Terraform, creating reusable, environment-specific modules for dev, test, and production deployments.
- Built and maintained CI/CD pipelines using Jenkins and GitHub Actions, enabling automated testing, validation, and controlled releases.
- Integrated NoSQL data stores (MongoDB, Azure Cosmos DB) to support semi-structured and event-driven workloads for analytics use cases.
- Implemented centralized logging, monitoring, and alerting using Azure Monitor and Application Insights, improving observability and reducing incident response time.
- Improved deployment reliability by resolving environmental drift and configuration inconsistencies, increasing release stability and reducing rollback scenarios.

Software Engineer, Stirtoni Software Pvt Ltd. - India.

Jun 2016 - Jun 2019

Product: Telecom Data Engineering & Real-Time Analytics Platform

Tools: Python, SQL, Apache Spark, Apache Airflow, ETL Tools, OLAP, OLTP, DDL, DML

- Designed and developed scalable ETL pipelines using Java and Python within the Apache Spark and Hadoop ecosystem, processing high-volume telecom CDR and network event data.
- Built distributed data processing workflows using Apache Spark (JVM-based), leveraging Spark SQL and DataFrame APIs for near real-time analytics and KPI generation.
- Implemented partitioned data ingestion strategies and optimized Spark execution (partitioning, parallelism, batch sizing) to resolve delayed CDR processing, reducing end-to-end pipeline latency by ~50%.
- Engineered DDL/DML SQL workflows to support OLTP-to-OLAP data synchronization, incremental batch loads, and schema evolution for analytical reporting systems.
- Developed data validation and quality checks (deduplication, null handling, schema enforcement) within Java/PySpark processing layers, ensuring accuracy and consistency of telecom KPIs.
- Created and maintained Apache Airflow DAGs with parameterized configurations, retries, and dependency management to orchestrate Java- and Spark-based batch workflows across environments.
- Collaborated with telecom domain experts to translate business use cases (churn analysis, usage trends, network performance metrics) into Java/Spark-driven analytical datasets.

PROJECT EXPERIENCE

Olympic Data Analytics using Azure

Tools: SQL, Python, Tableau, Spotfire, Excel, Git, GitHub

- Reduced execution time by 45% through SQL analysis on transformed data, optimizing query performance using Azure Synapse Analytics and SQL.
- Enabled processing of large datasets up to 50% faster by optimizing data transformations using Apache Spark within Azure Databricks and integrating with REST APIs.
- Reduced unauthorized access incidents by 100% by enhancing data security and governance through implementing Azure Role-Based Access Control (RBAC) and Azure Active Directory.
- Encountered reporting delays caused by inefficient data processing and addressed them by optimizing transformations with Apache Spark in Azure Databricks and integrating Power BI dashboards, resulting in a 50% increase in processing speed and accelerated data-driven decision-making.

Cloud Computing & Big Data Projects

Tools: Hadoop, Java, MapReduce, Maven, Shell Scripting, Text Processing.

- Improved data processing efficiency by 50% by designing Hadoop MapReduce jobs for large-scale social graph analysis and follower count aggregation.
- Reduced computation time by 40% through distributed matrix multiplication using custom Java Writable classes and optimized MapReduce workflows.
- Boosted ETL throughput and data quality by 45% by implementing customer data cleansing and transformation pipelines.
- Enhanced performance of graph partitioning and clustering tasks by 55% using Maven-packaged Java applications executed in local and distributed environments.
- Improved large-scale graph processing by implementing optimized Hadoop MapReduce workflows with custom Java classes, boosting data processing speed by 50%.

EDUCATION

THE UNIVERSITY OF TEXAS ARLINGTON

May 2025

Master of Science (M.S.), Major: Computer Science and Information Systems, Arlington, TX

Course work: Design and Analysis of Algorithms, Data Structures, Data Analysis & Managing Techniques, Artificial Intelligence, Web Data Management, Data Mining, Cloud Computing & Big Data, Machine Learning, Software Engineering-II, Numerical Methods.

PROJECTS

VOLUGRAD | Tools: React, Laravel, Node.js, SQL.

- Led a team to Build a volunteer management platform with task tracking, professor dashboards, real-time chat, and security features like 2FA and Captcha. Optimized for cloud deployment, responsive UI, and PWA support.

Facial Insight Attendance Management System | Tools: OpenCV, TensorFlow, Flask, MySQL, HTML5, CSS3, JavaScript.

- AI-driven attendance system using OpenCV, TensorFlow, Flask, and MySQL for automated, accurate tracking. Features facial detection, recognition, secure database management, and user-friendly UI. Reduced manual efforts and improved efficiency.

Image Captioning in Machine Learning | Tools: CNN, RNN, LSTMs, Transformers, TensorFlow, PyTorch, MS COCO, Flickr30k.

- Designed an AI-driven image captioning system integrating CNNs for visual feature extraction RNNs/Transformers for Text generation, Enhanced accuracy using attention mechanisms & evaluated performance on datasets for improved caption coherence.

Real-Time Ride Fare Estimation | Tools: Apache Spark, PySpark, Kafka, HDFS, Python

- Built a Spark-based system to estimate ride fares in real time by processing streaming trip data from Kafka and storing outputs in HDFS. Demonstrated ability to work with distributed data processing, real-time analytics, and big data ecosystem tools.

Audience Tracking using URL Shortener | Tools: Node.js, Express.js, MongoDB, HTML5, CSS3, Bootstrap.

- Built a URL shortener with real-time analytics, geographic tracking, and referral insights to enhance data-driven marketing and user engagement.

Block-Level Data Deduplication for Optimized Storage | Tools: AWS, PyCharm, MySQL Workbench, Chunking Algorithm, SHA 256.

- This project tackles the challenges of storage inefficiency and resource wastage due to redundant data in the big data era. Block-level data deduplication seeks to enhance storage efficiency and management by eliminating duplicate data blocks within storage systems. It also seamlessly uploads deduplicated data to the cloud for enhanced accessibility and backup.