

Pranav Kuramkote Sudhir

+18574659377 | pranavkuramkotesudhir@gmail.com | <https://www.linkedin.com/in/pranav-kuramkote-sudhir-a46359168/> | <https://github.com/pranavkuramkotesudhir>

Summary

Data Scientist with advanced expertise in data analytics tools and technologies, positioned between mid and senior-level experience. Noteworthy accomplishments include:- Developed a data ingestion portal, enhancing data quality for semiconductor processes by reducing entry errors by 95%.- Implemented scalable ETL processes, achieving a 40% reduction in processing times and substantial cloud cost savings.- Engineered machine learning solutions to optimize manufacturing workflows, decreasing material waste by 30% and attrition. Aims to leverage ability to quickly adapt and streamline data solutions enabling consumer banking data management projects.

Skills

Languages: Python (Expert), SQL, JavaScript, HTML

Services: AWS (EC2, S3, RDS, Lambda, Redshift), Apache Spark, Kafka, HDFS, Redis, PostgreSQL, MongoDB, Cassandra, Django, Flask, scikit-learn, TensorFlow, Power BI, Tableau

DevOps: Docker, Kubernetes, Jenkins, Git

Methodologies: Six Sigma, Lean Methodologies, Agile Methodologies

Data Management: Data Analytics, Data Quality Management, Data Visualization, Data Governance, Data Lineage

Work Experience

Veeco Instruments

San Jose, CA.

Data Scientist

Jul 2024 - Present

- Developed and deployed an end-to-end customizable data ingestion portal using JavaScript and Flask, incorporating CI/CD pipeline with Jenkins
- Implemented high-availability architecture using Docker containers and Nginx load balancing, achieving zero-downtime deployments through rolling updates
- Designed automated deployment system with GitHub webhooks for continuous integration and real-time code synchronization
- Engineered a dynamic dashboard that parses passdown run logs to generate executive summaries of tool status
- Developed data visualization interface using JavaScript to transform complex operational data into actionable insights
- Created automated backup solution to securely transfer tool computer files to central servers for data mining purposes
- Streamlined data preservation workflow to enable further analysis and insights extraction

HCL Technologies

Bangalore, India.

Member Technical Staff

Feb 2021 - Jun 2023

- Architected robust, fault-tolerant data pipelines using AWS ECS, capable of processing over 500GB daily with seamless data quality management and lineage observation, significantly lowering data reprocessing occurrences.
- Utilized serverless data processing methodologies through AWS Step Functions and Lambda to enhance scalability and introduced Six Sigma principles in operations to optimize workflows.
- Spearheaded the migration of on-premise data warehouses to AWS Redshift, tripling query performance and facilitating upwards scaling to better handle expanding data volume challenges, emphasizing data quality and agility methodologies.
- Developed a real-time dashboard using Flask, deployed it using Docker, and Jenkins CI/CD pipeline for automated deployment. Integrated SQLite authentication and Nginx reverse proxy to ensure secure user access and optimal performance.
- Improved RDS PostgreSQL database queries and implemented AWS DynamoDB for high-throughput operations, exhibiting a 70% reduction in query times and streaming efficiency.

Quest global Engineering

Bangalore, India.

Trainee Engineer

Nov 2019 - Oct 2020

- Engineered an automated data cleaning and preprocessing system using AWS Glue and S3, saving the team 10 hours per week and increasing data scientist productivity by 25%.
- Optimized Amazon Redshift queries for large-scale data retrieval, reducing average query execution time by 40% and improving overall system performance.
- Created comprehensive Power BI dashboards analyzing customer service KPIs, visualizing key metrics including regional service distribution, part frequency analysis, and performance trends.

Education

Northeastern University, Boston, MA

Master of Science, Data Analytics

May 2025

- Achievements: GPA: 3.89/4.0
- Courses: Data mining, MLOps, Database Management, Data Visualization

Project

News analysis pipeline

- Stock market news API with an AWS-based data pipeline orchestrated by Airflow and containerized using Docker.

Most probable character prediction

- A machine learning model for next-character prediction using TensorFlow, enhancing text generation capabilities.

Certifications

IBM Data science professional certification