

# RONITH REDDY PRODDUTURI

Santa Clara, CA | +1 629-273-8318 | [ronith0306@gmail.com](mailto:ronith0306@gmail.com) | [LinkedIn](#)

---

## PROFESSIONAL SUMMARY

Data Engineer and Machine Learning Engineer with 3+ years of experience building scalable data pipelines, real-time analytics systems, and AI-powered applications. Skilled in Python, PySpark, SQL, Apache Spark, and AWS, with expertise in ETL/ELT workflows, ML model deployment, and data visualization. Proven track record in processing 15TB+ of data daily, improving fraud detection accuracy by 30%, and reducing model deployment cycles by 40% through automation.

---

## TECHNICAL SKILLS

**Programming & Frameworks:** Python, SQL, PySpark, Apache Spark, Django, Flask, FastAPI, JavaScript, TypeScript, React.js, Node.js, C++, HTML5, CSS3, Bootstrap.

**Data Engineering & Big Data:** ETL/ELT Pipelines, Data Lakes, Data Warehousing, Snowflake, Amazon Redshift, Google BigQuery, Apache Airflow, AWS Glue, Apache Kafka, Real-Time & Batch Data Processing.

**Databases:** PostgreSQL, MySQL, MongoDB, DynamoDB, Redis, Oracle, SQL Server, NoSQL Databases.

**Cloud & DevOps:** AWS (ECS, Lambda, S3, CloudFront, API Gateway, SageMaker), Azure DevOps, Docker, Kubernetes, Helm, Terraform, Jenkins, Ansible, GitHub Actions, Infrastructure as Code (IaC), CI/CD Pipelines.

**Machine Learning & MLOps:** scikit-learn, TensorFlow, PyTorch, Amazon SageMaker, Model Deployment, Feature Engineering, Model Monitoring, A/B Testing, Experiment Tracking (MLflow).

**Analytics & Visualization:** Pandas, NumPy, Tableau, Power BI, Matplotlib, Seaborn, Plotly, KPI Dashboards, Statistical Analysis, Data Storytelling.

**Testing & Quality Assurance:** pytest, Selenium, Unit Testing, Integration Testing, TDD/BDD, Debugging, Code Review.

**Methodologies & Architecture:** Agile/Scrum, Kanban, REST API Design, OOP, MVC/MVT Architecture, Microservices Architecture, Event-Driven Architecture.

---

## PROFESSIONAL EXPERIENCE

### Full Stack Python Developer

Mar 2025 - Present

#### PayPal | San Jose CA

- Designed and deployed Flask-based microservices leveraging SQLAlchemy and OOP patterns, sustaining 99.9% uptime while supporting high-volume data transactions.
- Automated CI/CD pipelines in Azure DevOps with Ansible, integrating pytest and functional testing to reduce deployment issues by 40%.
- Tuned SQL and NoSQL queries monitored through Azure App Insights, leading to a 50% cut in API and dashboard latency for end users.
- Delivered real-time KPI dashboards in Tableau sourced from Spark-driven ETL jobs, transforming multi-hour reporting processes into on-demand analytics.
- Built Python-powered data validation and cleansing workflows, improving ETL dataset accuracy by 35% for downstream analytics.
- Optimized Apache Spark job configurations for parallelism, boosting batch data processing speed by over 40% in ML feature generation tasks.
- Presented query optimization metrics and data throughput insights to cross-functional teams, prompting infrastructure adjustments that increased efficiency by 25%.
- Produced clear documentation outlining ML data workflows, enabling faster onboarding and ensuring consistency in analytics processes.

### Full Stack Python Developer

Apr 2024 - Feb 2025

#### Vanguard | Remote

- Delivered FastAPI and GraphQL fraud detection services that processed thousands of concurrent user requests, cutting confirmed fraud incidents by 30% through real-time ML scoring.
- Reduced model training preparation time by 45% by migrating fraud datasets to Snowflake data lakes with AWS Glue ETL workflows, streamlining the ML pipeline.
- Achieved sub-300ms prediction latency for anomaly detection by deploying AWS Lambda applications backed by DynamoDB and containerized inference endpoints.
- Boosted analyst productivity by 25% by developing React.js dashboards that visualized streaming ML outputs from Python APIs for instant risk assessment.
- Maintained 24/7 service availability by orchestrating Kubernetes-based deployments via Jenkins and GitHub Actions, eliminating production downtime during releases.

- Increased stakeholder trust in ML predictions by implementing SHAP and LIME explainability, enabling clear insight into model decision-making.
- Elevated deployment quality by coaching junior engineers on Terraform automation, AWS Lambda optimization, and secure API design, leading to a 20% improvement in release success.
- Ensured reproducible ML model performance by establishing MLflow-based version control and experiment tracking in collaboration with data science teams.

## Python Developer

Oct 2020 - Dec 2022

### Bayer | India

- Delivered optimized React.js and Node.js front-end modules integrated with Django and Flask backends over GraphQL, enabling customers to complete transactions 20% faster through reduced API response times.
- Architected a Django-powered analytics framework with ML-driven insights, leveraging AWS ECS and Lambda to handle over 15TB of daily transaction data, resulting in a 35% increase in throughput.
- Orchestrated real-time fraud detection by deploying Dockerized Flask microservices with SQLAlchemy to Kubernetes, achieving sub-second predictions via SageMaker and ensuring seamless scalability.
- Transformed multi-source transactional data pipelines using PySpark to load into Snowflake data lakes, cutting batch latency by 40% and making ML features available ahead of schedule for model training.
- Enhanced deployment reliability by implementing Jenkins and Ansible pipelines with PyTest validation, which reduced production release issues by 25% and improved post-deployment stability.
- Improved schema consistency and indexing for MySQL and Redis ETL jobs, raising prediction accuracy of downstream ML models by 18% and eliminating recurring data integration errors.
- Led a cross-functional team of five to operationalize ML models with SageMaker and MLflow tracking, reducing deployment cycles by 40% and maintaining reproducibility across all releases.
- Designed monitoring dashboards with Prometheus and Grafana to track API health and latency, sustaining 99.99% uptime for core payment services and enabling proactive incident handling.

## PROJECTS

### Fraud Detection & Prevention System

- Developed PySpark and Snowflake workflows with scikit-learn models to scan 10TB+ transaction data daily, lowering fraudulent activity by 30% through predictive scoring.
- Automated AWS SageMaker deployments with CI/CD pipelines, shortening release cycles from two weeks to three days while ensuring model reproducibility.
- Connected FastAPI microservices to live Power BI dashboards, giving fraud teams real-time insight into transaction anomalies and risk scores.

### IoT Health Monitoring Platform

- Implemented AWS Kinesis and Lambda pipelines for ingesting 1M+ sensor events daily, ensuring low-latency data flow to hospital analytics systems.
- Applied Python-based anomaly detection algorithms to identify early signs of health deterioration, achieving 92% classification accuracy.
- Built React.js interfaces displaying live patient vitals, enabling medical teams to act on alerts instantly.

### Real-Time Language Translator AI App

- Created a multilingual translation engine with Whisper API, Flask, and React Native, supporting 12 languages under 300ms latency for spoken text.
- Set up automated PyTest pipelines to maintain translation accuracy within 0.2% deviation during updates.
- Deployed AWS Lambda and API Gateway backend for scalable, cost-efficient real-time processing.

## EDUCATION

### Master of Science in Technology Management

Aug 2023-Dec 2024

Lindsey Wilson University | KY

### Bachelor of Technology in Computer Science

Jun 2018-Jun 2022

GITAM University | India

## CERTIFICATIONS

- AWS Certified Data Engineer – **Associate**
- AWS Certified Solutions Architect - **Associate**