

SRINIVASA REDDY KOMMIREDDY

Generative AI Engineer & Data Engineer | Python · SQL · Snowflake · Databricks · GenAI

+1 805-871-6548 | srinikreddy1998@gmail.com | [LinkedIn](#)

PROFESSIONAL SUMMARY

- Generative AI Engineer and Data Engineer with 4+ years of experience building real-time data pipelines and production Generative AI systems across large-scale healthcare and financial claims platforms.
- Hands-on experience integrating Generative AI into healthcare claims workflows - OpenAI API-based denial-explanation assistants, Retrieval-Augmented Generation (RAG) over coding and policy documentation, and prompt engineering - to accelerate root-cause analysis and reduce manual case-review time.
- Experienced in SQL, Python, Snowflake, Databricks, and Power BI, analyzing 50M+ claims and \$150M+ in healthcare spend to improve revenue-cycle performance, claims operations, and financial visibility.
- Proven track record identifying \$4.2M in reimbursement recovery opportunities, reducing avoidable claim denials by 15%, and building scalable analytics and GenAI-enabled solutions that enable data-driven decision-making across Finance and Operations.
- Built and maintained HIPAA-compliant data pipelines in Azure Databricks (PySpark) to ingest and standardize EDI 837/835 and FHIR datasets, achieving 99.9% data integrity for audit readiness.
- Engineered feature datasets in Python (pandas, scikit-learn) for claim-denial propensity modeling, partnering with data science teams to support model development, monitoring, and GenAI-based output explanation.
- Experience architecting claims analytics data models (Databricks, SQL, PySpark) centralizing policy terms, trigger attributes, and claims data as a single source of truth for pre-approval validation and operational reporting.
- Skilled in Power BI semantic modeling (DAX, drill-through, row-level security), enabling 25-35% safe auto-approval rates and reducing median claims triage time by 87%.
- Applied statistical hypothesis testing and experiment design (A/B testing) to validate process improvements, confirming a \$2M reduction in manual rework costs.
- Explored early Generative AI applications (prompt engineering, OpenAI API) for claims-triage rationale summarization, laying the groundwork for production GenAI adoption in subsequent roles.
- Microsoft Certified: Fabric Analytics Engineer Associate and Power BI Data Analyst Associate, with a strong cloud data engineering background across Azure and AWS.
- Collaborated cross-functionally with Finance, Operations, Actuarial, and Claims teams to translate business and regulatory requirements into governed, production-grade analytics and AI systems.

TECHNICAL SKILL MATRIX

Generative AI & LLMs (Applied)

OpenAI API, Prompt Engineering, Retrieval-Augmented Generation (RAG), Azure OpenAI, Vector Search / Semantic Search, LLM Output Evaluation

Programming

Python (Pandas, NumPy, PySpark, Scikit-learn), SQL (CTEs, Window Functions, Stored Procedures), Data Validation, Feature Engineering, Automation, API Integration

Data Engineering & Modeling

Advanced SQL (Performance Tuning), Data Modeling (Star, Snowflake, Dimensional), Real-Time & Streaming Pipelines (Spark Structured Streaming), Apache Airflow, Incremental & CDC Pipelines, Data Quality Testing, Schema Management, ETL/ELT

Analytics & Modeling

Revenue Cycle Analytics, Cohort & Funnel Analysis, Time-Series Analysis, Experiment Design (A/B Testing), Churn & Retention Modeling, KPI Frameworks, Root Cause Analysis, ROI & Forecasting

Cloud Technologies

Azure (ADLS Gen2, Data Factory, Databricks, Synapse, Azure OpenAI), AWS (S3, Glue, Lambda, Redshift, Kinesis), Data Lakes, Serverless Pipelines, Cost Optimization, Security & RBAC

Data Visualization & Reporting

Power BI, Tableau, Advanced Excel (Power Query, Pivot Tables, VBA Macros), Looker, SSRS, Data Storytelling, Dashboard Design, KPI Reporting

Databases & Data Warehousing

Snowflake, PostgreSQL, MySQL, MSSQL, SQL Server, BigQuery

Certifications

Microsoft Certified: Fabric Analytics Engineer Associate, Microsoft Certified: Power BI Data Analyst Associate

PROFESSIONAL EXPERIENCE

UnitedHealth Group | February 2025 - Present

Generative AI Engineer & Data Engineer

Responsibilities:

- Architected and engineered a GenAI-powered denial-explanation pipeline using the OpenAI API, prompt engineering, and Python microservices, auto-generating plain-language explanations of claim-denial-propensity model outputs in near real time for claims analysts, reducing manual case-review time.
- Built a RAG-based coding and policy knowledge engine using Azure OpenAI, vector embeddings, and semantic search over CMS coding guidelines and payer policy documentation, exposing a queryable API layer that enabled natural-language retrieval of complex ICD-10/CPT coding rules and accelerated root-cause coding-gap analysis.
- Engineered real-time and batch data pipelines in Azure Databricks (PySpark) processing 50M+ claims and \$150M+ in healthcare spend, powering Medical Loss Ratio and AR-aging metrics with near-real-time data freshness for monthly close cycles.
- Designed and built SQL-driven ETL/ELT pipelines performing large-scale root-cause exploration across ICD-10, CPT/HCPCS, and authorization records; identified coding gaps that reduced avoidable denials by 15%, driving an estimated \$4.2M increase in annual reimbursement recovery.
- Built and maintained feature-engineering pipelines in Python (pandas, scikit-learn, PySpark) for claim-denial propensity modeling, incorporating provider behavior, historical denial patterns, and utilization signals into production-ready feature stores for data science model training.
- Applied prompt-engineering and evaluation practices to validate AI-generated denial explanations against clinical and coding accuracy standards, ensuring HIPAA-compliant, audit-ready outputs at production scale.
- Engineered a statistical validation pipeline in Python to run hypothesis-driven, significance-tested analysis of processing trends at scale, confirming denial reductions were driven by process improvements rather than seasonal volume shifts, validating a \$2M reduction in manual rework costs.
- Established HIPAA-compliant, production-grade data pipelines in Azure Databricks (PySpark) to ingest and standardize EDI 837/835 and FHIR datasets in near real time, implementing automated data-quality checks that achieved 99.9% data integrity for audit readiness.
- Engineered the underlying Snowflake data model and real-time pipeline powering Power BI dashboards used to track Medical Loss Ratio and AR aging, and partnered with Finance and Operations to optimize authorization-workflow pipelines, reducing processing delays by 22%.

Environment: OpenAI API, Azure OpenAI, RAG, Prompt Engineering, Vector Search, Python, SQL, Snowflake, Azure Databricks, PySpark, Scikit-learn, Power BI, EDI 837/835, FHIR

Data Engineer

Responsibilities:

- Architected a real-time Databricks claims analytics data model using SQL and PySpark to centralize policy terms, trigger attributes, and claims data, establishing standardized parametric and indemnity claims logic as the single real-time source of truth for pre-approval validation and operational reporting.
- Explored early Generative AI applications (OpenAI API, prompt engineering) to prototype natural-language summarization of claims-triage rationale and policy-exception explanations, laying groundwork for broader GenAI adoption within the claims data engineering team.
- Engineered rule-based logic-gate pipelines and exception-reporting workflows (SQL, PySpark, DAX) linking claims outcomes with policy and trigger data in near real time, enabling 25-35% safe auto-approval while reducing median claims triage time by 87% from 24 hours to 3 hours.
- Built automated, real-time data-quality and reconciliation pipelines in Databricks using PySpark and SQL to enforce schema consistency, null and range validation, and cross-field business rules, reducing manual validation effort by 50% and reconciliation exceptions by 40%.
- Engineered the semantic data layer and row-level-security model powering Power BI operational dashboards used to monitor claims SLAs, auto-approval performance, and exception drivers in real time across actuarial and claims teams.
- Partnered with Actuaries and Claims Leads to translate evolving policy language and risk thresholds into governed, production SQL and PySpark business rules, maintaining controlled logic updates and metric-consistent reporting while accelerating low-complexity claims payouts.

Environment: OpenAI API (Exploratory), Prompt Engineering, SQL, PySpark, Databricks, Real-Time Pipelines, Data Quality, Reconciliation Pipelines, Power BI, DAX

EDUCATION

Master of Science in Information Technology - St. Francis College | Aug 2023 - Dec 2024

CERTIFICATIONS

- Microsoft Certified: Fabric Analytics Engineer Associate
- Microsoft Certified: Power BI Data Analyst Associate