Nagendra Shivasai Kanneboina
Intern Data Engineer
Location: Ashburn, VA | US Citizen
Contact: (952)-994-0235 | Nagendrashivasai.kanneboina@gmail.com

SUMMARY

Experienced in building secure systems, conversational AI agents, and RAG-based search and summarization solutions. Developed cloud-connected applications and full-stack features using Python, JavaScript, React, and Node.js. Worked with vector databases, LLM APIs, and ML models for forecasting, reporting, and content delivery. Supported large-scale data engineering initiatives, including Databricks migration, automated SQL validation, ETL optimization, and cloud-based pipeline development using Azure ADF across Oracle and Teradata systems. Analyzed legacy IBM DataStage jobs and designed equivalent ADF pipelines, data flows, and orchestration logic while documenting transformations, source-to-target mappings, and data processes to support data consistency and enterprise standards. Performed detailed reconciliation and data validation to ensure accuracy and alignment between DataStage and ADF outputs. Built machine-learning frameworks and Power BI dashboards for anomaly detection and operational reporting, helping identify financial irregularities and support data-driven decision-making through research, analysis, and cross-functional collaboration.

EDUCATION

Bachelor of Science in Computer Science – George Mason University 2023 – exp* 2027

PROFFESIONAL EXPERIENCE

ETL Intern – Prak Technologies | September 2025 – December 2025

Maintained and optimized Oracle and Teradata databases containing ~700k Medicare & Medicaid member records during migration from legacy platforms to enterprise cloud systems. Supported data analysis, validation, and reconciliation efforts to ensure data accuracy, consistency, and alignment across systems during Databricks migration and UAT activities. Assisted with ETL workflows, automated SQL validation, and troubleshooting data integrity issues using Azure ADF while documenting pipelines, mappings, and data processes following enterprise standards. Analyzed existing IBM DataStage jobs, mappings, and cross-system dependencies to support migration readiness and pipeline development. Designed and implemented equivalent Azure ADF pipelines, data flows, and orchestration logic, rewriting complex transformations using ADF expressions, SQL, and Databricks. Executed row-level and column-level reconciliation, documented validation results, resolved data mismatches, and collaborated with cross-functional teams to support reporting, production cutover, and ongoing data operations.

PROJECTS & HACKATHONS

AI-Driven Anomaly Detection for Medicare Broker Commission Validation – BCBSRI Broker

Developed a machine learning framework to analyze Medicare broker commission payments and identify irregularities across enrollment, hierarchy, product, and payout data. Built an end-to-end data analysis and anomaly detection pipeline using Python, SQL, and Scikit-learn, leveraging Isolation Forest and feature engineering to support data validation, accuracy, and risk analysis. Implemented automated risk scoring processes to streamline manual validation tasks and improve operational efficiency. Designed and maintained Power BI dashboards and reports to visualize anomaly trends, risk indicators, KPI metrics, and commission variance patterns for business analysis and decision-making. Supported documentation of data processes, validation logic, and analytical workflows while collaborating across teams to enhance reporting, financial governance, and data-driven operations.

Machine-learning-powered International Economic Data Prediction - Hop Hacks

Built an ML-driven analytics platform that processes global economic data to generate forecasts, insights, and data-driven recommendations. Performed quantitative analysis on historical datasets and developed time-series forecasting models for GDP, inflation, trade, and market stability metrics. Implemented RAG pipelines to improve reliability and consistency of LLM-generated outputs using validated enterprise datasets.
Designed cloud-based data ingestion and transformation pipelines to support automated data processing, continuous model updates, and scalable deployment. Developed dashboards, reports, and chatbot/voice bot interfaces to deliver analytics summaries and interactive insights, while supporting documentation, data mapping, and cross-functional collaboration throughout the project lifecycle.

AI powered Note Stream - Patriot Hacks

Built an intelligent note curation platform that aggregates, organizes, and summarizes user-generated and external content into a centralized knowledge hub. Performed data analysis and content processing to improve information retrieval, personalization, and recommendation accuracy. Implemented AI-driven features to generate adaptive learning paths, summaries, and search capabilities using vector databases, RAG pipelines, and LLM APIs. Developed and maintained full-stack and cloud-based features using Python, JavaScript, React, Node.js, and cloud platforms including AWS, Azure, and GCP. Designed scalable data ingestion and processing workflows, supported documentation of system components and data flows, and collaborated across project teams to enhance platform performance, reliability, and user experience.

TECHNICAL SKILLS

Languages & DBs: Python, Java, C, C++, SQL, HTML, CSS, JavaScript, Oracle, Snowflake DB, SQL Server, Azure DB, Teradata, DBT
Frameworks & Tools: Pandas, NumPy, Scikit-learn, UNIX commands/scripting, Autosys, GitHub, Mathlib, Microsoft Excel, Jupyter Notebook
Cloud & Analytics: AWS (S3, Route 53, Amplify), Azure ADF, PySpark, Power BI, RAG pipelines, OpenAI APIs, Microsoft Copilot Studio
Data & Processing: SQL-based data simulation/extraction, data preprocessing, feature engineering, dashboarding/visualization

CERTIFICATIONS

Foundry & AIP Builder Foundations – Palantir Technologies 2025 Machine Learning Foundations – AWS 2025