

# POORNA SAI VEERA HARISH PEDAPUDI

Virginia (Open to Relocation) | +1 (571) 281-1394 | [poornasaiapedapudi@gmail.com](mailto:poornasaiapedapudi@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## SUMMARY

Data Analyst with 5+ years of experience delivering insights from 1B+ records across finance, education, and SaaS sectors. Spearheaded analytics pipelines and predictive models that improved reporting speed by 30% and lifted campaign conversion by 15%. Proficient in Python, SQL, Tableau, AWS, and ML tools, having optimized ETL workflows and deploying end-to-end solutions that influence strategic decisions.

## TECHNICAL SKILLS

**Languages & Tools:** Python, SQL, R, Bash, Excel, Power BI, Tableau, Jupyter Notebook, Git.

**Libraries & Frameworks:** pandas, numPy, scikit-learn, matplotlib, seaborn, PySpark, XGBoost, LightGBM, PyTorch, LangChain.

**Cloud & Databases:** AWS (S3, Redshift, Glue, Athena), Snowflake, MySQL, PostgreSQL, MongoDB, BigQuery, Databricks.

**Analytics & Machine Learning:** EDA, A/B Testing, KPI Design, Time Series Forecasting, Predictive Modeling, Statistical Testing.

**Core Concepts:** Pipeline Automation, Dashboard Design, SQL Optimization, Feature Engineering, Model Evaluation.

**Other Skills:** API Integration, Version Control (Git/GitHub), Prompt Engineering, CI/CD, Experiment Tracking (MLflow), Data Ethics.

## PROFESSIONAL EXPERIENCE

Indukuri Information Services Inc, (Remote)

Oct 2023 – Present

**AI/ML Engineer (Volunteer)**

Mar 2024 – Present

- Performed exploratory trend modeling and A/B testing using scikit-learn and stats models, identifying data-driven optimizations in marketing strategies.
- Streamlined early-stage ETL workflows and basic data pipelines using Python, PySpark, SQL, and Apache Airflow, enhancing operational efficiency and supporting delivery of financial reporting insights.

**Data Analyst (Volunteer)**

Oct 2023 – Feb 2024

- Architected initial database structures using SQL and AWS S3, laying the foundation for scalable reporting, automated querying, and improved data consistency.
- Constructed SQL-based analytics pipelines integrated with AWS Redshift, accelerating data retrieval by 25% and improving pilot reporting projects.

**Data Analyst at Hildreth Institute (Fairfax, VA)**

Jan 2024 – May 2024

- Evaluated 250K+ education records to build KPIs and predictive Tableau dashboards, uncovering trends that supported education policy decisions.
- Visualized complex financial data by designing 10 Tableau dashboards; enhanced data communication by 35% and facilitated clearer decision-making regarding cost of attendance metrics for stakeholders.
- Simplified data extraction and ingestion using Python (APIs, web scraping) and SQL, reducing manual data entry by 90% and optimizing data storage with AWS S3 and Redshift.
- Architected AWS S3 and Redshift-based infrastructure to manage 100+ GB of data, boosting retrieval speed by 40%.

**Data Scientist at Tritech Software Pvt Ltd, (Hyderabad, India)**

Jul 2019 – Dec 2022

- Orchestrated the migration of data from legacy ERP systems into a centralized data warehouse using Python and SQL, diminishing data silos and accelerating reporting cycles by 30%.
- Leveraged EDA techniques to uncover seasonal demand trends, optimize inventory forecasting, reduce stockouts by 25%, and enhance supply chain planning accuracy.
- Built Tableau dashboards to visualize KPIs like revenue growth and regional performance, reducing reporting time by 30% and improving stakeholder insights.
- Revamped ETL workflows and unified diverse sources, boosting predictive model performance by 20% and reducing inference time by 30%.

## PROJECTS

**NLP-Based Review Analyzer & Price Tracking System**

Python, BeautifulSoup, AWS Lambda, S3, SQL, MongoDB, LangChain, Prompt Engineering, Data Ethics

- Developed a real-time scraper and AWS Lambda pipeline to capture 10K+ product prices/min, with review analysis using LangChain and OpenAI for sentiment extraction.
- Ensured ethical AI practices by obfuscating review metadata and integrating prompt engineering techniques to extract actionable product insights.

**Customer Behavior Analysis & Campaign Optimization**

Python, R, SQL, Scikit-learn, Tableau, A/B Testing, XGBoost, MLflow, CI/CD Concepts

- Built predictive models using XGBoost to segment 100K+ bank customers, improving campaign targeting and increasing conversions by 15%.
- Tracked model performance using MLflow and implemented version-controlled pipelines aligned with CI/CD workflows for deployment readiness.

## EDUCATION

**MS in Data Analytics Engineering** at George Mason University, Fairfax, VA

Jan 2023 – May 2024

**BS in Electronics and Communication Engineering** at Malaviya National Institute of Technology, Jaipur, India

Jul 2016 – May 2020