

Gouthami Nadupuri

+1(341)-356-3732 | gouthaminadupuri25@gmail.com | [LinkedIn](#) | [GitHub](#)

EDUCATION

University of the Pacific
Master of Science in Data Science

Aug 2023 – May 2025
San Francisco, CA

Gitam University
Bachelor of Technology in Computer Science and Engineering

Jul 2008 – May 2012
Vizag, India

EXPERIENCE

Data Quality Analyst

Jan 2020 – June 2023

Vitech Systems

Hyderabad, India

- **Project:** National Benefit Trust Fund
- Extracted and validated large datasets related to pension and insurance plans for analysis and reporting.
- Conducted data quality checks, statistical analysis, and root cause investigations to ensure accuracy.
- Developed SQL queries and Excel models to analyze member data, track claims, and generate reports.
- Built dashboards to provide real-time visibility into claims processing and member benefits.
- **Project:** Motion Picture Industry Pension & Health Insurance
- Conducted data quality analysis on pension and healthcare datasets to identify anomalies and compliance issues.
- Automated data validation and reporting tasks using SQL and Python scripts.
- Created dashboards and visualizations to present insights to client stakeholders.
- Partnered with business users to convert requirements into analytical reports.
- Recognized for improving accuracy and reliability of business-critical data.
- **Tools/Skills:** SQL, Python (Pandas, NumPy), Tableau, Power BI, Data Quality, Data Visualization, Data Automation

Senior Data Analyst

Mar 2018 – Nov 2019

Wipro

Hyderabad, India

- **Project:** Google Trust & Safety
- Conducted in-depth data analysis on e-commerce merchant activities to ensure compliance with Google Shopping policies.
- Designed and maintained data quality reports, identifying violations and providing actionable insights.
- Automated recurring reporting workflows, reducing manual effort by streamlining data processes.
- Trained junior analysts in data interpretation, visualization, and reporting best practices.
- Delivered insights to improve customer trust and user experience.
- **Tools/Skills:** SQL, Excel, Tableau, Data Quality Analysis, Automation (Macros), Reporting

Data Analyst

Aug 2016 – Dec 2017

Accenture

Hyderabad, India

- **Project:** E-commerce
- Collected, validated, and analyzed e-commerce transaction data to support reporting and performance tracking.
- Used SQL queries to extract, clean, and transform raw data for analysis.
- Built Excel dashboards and visual reports to monitor customer purchase trends and product performance.
- Collaborated with stakeholders to identify data inconsistencies and recommend improvements.
- **Tools/Skills:** SQL, Excel, Data Cleaning, Data Validation, Business Reporting

PROJECTS

Customer Churn Prediction | Machine Learning

- Delivered retention strategy insights based on customer service usage and behavioral patterns, improving churn prediction accuracy.
- Created visualizations (Matplotlib, Seaborn) to communicate insights to non-technical stakeholders.
- **Skills/Tools:** Python, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, SMOTE, GridSearchCV

Fake News Detection | Natural Language Processing (NLP)

- Built an NLP pipeline (text preprocessing, tokenization, TF-IDF, word embeddings) for news classification.
- Implemented Logistic Regression, Random Forest, and BERT-tiny models for performance benchmarking.
- Achieved high accuracy and F1-score, with BERT-tiny outperforming classical models.
- Visualized model comparisons and misclassifications using Seaborn & Matplotlib.
- **Skills/Tools:** Python, NLTK, Scikit-learn, Transformers (Hugging Face), TensorFlow/PyTorch, Pandas, NumPy, Matplotlib, Seaborn

E-Commerce Fraud Detection System | Data Engineering

- Designed a real-time fraud detection pipeline using Apache Kafka for streaming transaction data.
- Integrated machine learning models for anomaly detection, risk scoring, and automated fraud alerting.
- Used Elasticsearch + Kibana dashboards for anomaly visualization and fraud monitoring.
- Implemented ETL workflows for ingesting, cleaning, and transforming high-volume transaction data.
- **Skills/Tools:** Apache Kafka, Python, Spark (PySpark), Machine Learning, Kibana, Elasticsearch, SQL

Time Series Forecasting | Data Science

- Built ARIMA and Prophet models to forecast sales with trend and seasonality decomposition.
- Conducted exploratory data analysis (EDA) and feature engineering to identify seasonality and outliers.
- Evaluated model performance using RMSE and MAPE, achieving improved forecasting accuracy.
- Developed visualizations and reports to support business planning and decision-making.
- **Skills/Tools:** Python, Stats models, Facebook Prophet, Pandas, NumPy, Matplotlib, Seaborn, Time Series Analysis

TECHNICAL SKILLS

Programming: Python, R, SQL, C++, FastAPI, HTML, CSS, JavaScript

Data Analysis & Visualization: Tableau, Power BI, Pandas, Numpy, Matplotlib, Seaborn, Plotly, Looker

Machine Learning & AI: Supervised and Unsupervised, Scikit-learn, TensorFlow, NLP, LLM, PyTorch, Deep Learning, Prompt Engineering, Precision/Recall, F1 Score, LangChain

Data Engineering & Cloud: ETL, Apache Kafka, NoSQL, AWS, CI/CD, Docker, Kubernetes

Tools & Platforms: Jupyter, Google Colab, VS Code, PyCharm, Git, Postman, Jira, Confluence, API

CREDENTIALS & VOLUNTEER CONTRIBUTIONS

- AWS Cloud Practitioner Essentials
- Python Programming Masterclass with Data Structures and Algorithms
- IBM RAG and Agentic AI
- Volunteer and active team member in Google Developers Group (GDG) and Hacker's quad tech communities, supporting event organization, knowledge-sharing sessions, and community engagement initiatives.