

Aditi Patil Raman

+1 (704) 369 4543 | aditi.rpatil2001@gmail.com | United States (Open to Relocation) | [linkedin/aditipatil](https://www.linkedin.com/in/aditipatil)

EDUCATION

Master of Science in Information Technology, GPA: 3.8 / 4.0 | University of North Carolina Charlotte

Coursework: Business Intelligence and Analytics, Knowledge Based Systems, Applied databases, Artificial Intelligence.

Aug 2023 – May 2025

Bachelor of Engineering in Computer Science, CGPA: 8.5 / 10 | Visvesvaraya Technological University

Coursework: Database Management Systems, Big Data Analytics, Data Mining, Visual Analytics, Machine Learning.

Jul 2019 – Jun 2023

TECHNICAL SKILLS

- Programming & Scripting:** Python, NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, SQL, R Programming
- AI & ML:** Regression & Classification Models, Ensemble Methods (Random Forest, XGBoost), Neural Networks (TensorFlow, PyTorch, LSTM), NLP & LLMs (LangChain, OpenAI API, RAG, Prompt Engineering), Time Series Forecasting, Feature Engineering, Model Evaluation, PyCaret, SHAP
- Data Engineering & Warehousing:** ETL Pipelines, Data Cleaning & Transformation, Data Modeling, Snowflake, PostgreSQL
- Big Data & Streaming:** Apache Kafka, Hadoop, Spark, AWS EMR
- Cloud Platforms & Services:** AWS (Redshift, Glue, S3, EC2, Lambda, Athena, IAM, CloudWatch, CloudFormation), Google Cloud Platform
- Visualization & Reporting:** Tableau, Power BI, Looker Studio, Amazon QuickSight, Streamlit
- Tools & Frameworks:** Flask, FAISS, Hugging Face, DBT, Apache Airflow, Jupyter Notebook, Git/GitHub, Docker

WORK EXPERIENCE

Graduate Assistant - Data Engineer | UNC Charlotte

Charlotte, NC | Oct 2023 – May 2025

Technologies: SQL, PostgreSQL database, AWS Glue, Amazon S3, Amazon Redshift, Tableau.

Pipeline Automation & Data Governance

- Engineered ETL pipelines using AWS Glue and S3, processing ~10,000 JSONs and automating schema cataloguing, reducing manual prep by 30%.
 - Implemented data validation checks and schema governance frameworks in AWS Glue/Redshift, ensuring data accuracy across 15+ departments.
 - Partnered with academic advisors to define KPIs, validate key metrics, and refine data models, enhancing the reliability of academic reporting.
- Analytics, Optimization & Insights
- Optimized Amazon Redshift schemas with distribution and sort keys, improving query performance and enabling faster cross-department analytics.
 - Constructed advanced SQL (CTEs, window functions) to extract and aggregate student, course, and grade data for multi-department dashboards.
 - Designed and deployed interactive Tableau dashboards integrating Redshift/PostgreSQL and Banner system data, transforming raw datasets into actionable insights and reducing manual tracking by 25%.

Data Engineer | Bharath Electronics Limited

Bangalore, India | Jun 2022 – Dec 2022

Technologies: SQL, Tableau, Python(Pandas), Jupyter Notebook

- Executed data transformation in Python (Pandas), performing data preprocessing, aligning schemas and filters to improve data consistency by 25%.
- Exported cleaned datasets to MySQL using SQLAlchemy for efficient fault detection and advanced component-level analysis.
- Wrote complex SQL queries (JOINS, CTEs, CASE) to identify recurring faults, anomalies, and high-risk components across datasets.
- Built Tableau/Power BI dashboards visualizing fault trends, error patterns, and KPIs for actionable insights, reducing manual analysis time by 30%.

PROJECTS

Health Insurance Cost Prediction using Scikit-learn ([GitHub](#)) | Python, Scikit-learn, Pandas, Seaborn, NumPy, Google Colab, Matplotlib.

- Trained a ML model to predict charges using one-hot encoding and exploratory data analysis with Seaborn and Matplotlib to identify key cost factors.
- Applied Linear Regression ($R^2 = 0.78$) and optimized performance using Polynomial Regression, improving R^2 to 0.86 and reducing MAE (4284 → 2599) and RMSE (5796 → 4570) through feature engineering and model tuning.

Real-Time Event Streaming Pipeline for E-Commerce | Apache Kafka, Python, AWS EC2, AWS Lambda, AWS Glue, S3, Athena.

- Orchestrated a Kafka pipeline on EC2 and Lambda to process user events and stored the data in S3, with Python simulating random events.
- Integrated Glue to catalog and structure data in S3 and ran Athena queries for analytics on customer behavior and product activity.

Predictive Employee Turnover Analysis ([GitHub](#)) | Google Cloud BigQuery, PyCaret, Random Forest, Looker Studio

- Set up a BigQuery database, connected Google Colab to develop a churn prediction model using PyCaret and Random Forest.
- Produced a visual report with employee satisfaction (50%) and performance (47%) for data-driven strategies.
- Analyzed turnover patterns and identified job satisfaction as the top driver of churn, along with factors like tenure and workload.

PredictWise – AI Sales Forecast ([GitHub](#)) | Python, Pandas, NumPy, Scikit-learn, PyTorch, XGBoost, Matplotlib, PySpark, Streamlit.

- Designed an AI-driven forecasting system to predict future sales and optimize promotional pricing using retail transaction data and temporal features.
- Trained and compared XGBoost and LSTM models against ARIMA and regression baselines, achieving a 12% lift in forecast precision and $R^2 = 0.94$.
- Created an interactive Streamlit dashboard for real-time sales prediction and scenario simulation, reducing manual reporting and analysis time by 78%.

GenQuery – LLM-Based SQL Interpreter ([GitHub](#)) | Python, Flask, LangChain, FAISS, Hugging Face, OpenAI GPT-4, Meta LLaMA-3.

- Developed an LLM-powered RAG pipeline using LangChain, FAISS, and Hugging Face embeddings, improving semantic query relevance by 92%.
- Integrated OpenAI GPT-4 and Meta LLaMA-3 for hybrid reasoning and fallback generation, boosting factual accuracy by 18%.
- Deployed a Flask-based app for natural-language SQL querying on IMDb data with <1.2 s latency and 40% lower memory use.

CERTIFICATIONS

- Machine Learning Specialization — Stanford University & DeepLearning.AI
- AWS Certified AI Practitioner [[Credly Link](#)]
- AWS Certified Solutions Architect – Associate [[Credly Link](#)]
- AWS Certified Data Engineer - Associate [[Credly Link](#)]
- Google Data Analytics Professional Certificate [[Credly Link](#)]

Oct 2025

Feb 2025

Feb 2025

Jul 2024

Feb 2024