

TAYOUTH MALLA

TX, USA | 940-629-9923 | tayouthmalla@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Data Analyst and Data Engineer with 4+ years of experience in **data analytics, SQL reporting, and scalable ETL pipelines**. Skilled at transforming complex datasets into actionable insights that drive business outcomes. Experienced in **cross-functional collaboration** with engineering, product, and compliance teams to improve decision-making and developer experience. Pursuing a Master's in Advanced Data Analytics (GPA 4.0) with a strong passion for **apps, developer ecosystems, and process improvement**.

WORK EXPERIENCE

Verity Advisors | Data Engineering Intern

Feb 2025 - May 2025

- Designed **ETL pipelines** (AWS Glue, Lambda, Redshift) to process **1M+ records weekly**, reducing reporting latency by 40%.
- Built automated reporting workflows with **real-time stock data ingestion**, API integrations, and AWS SES for **daily stakeholder updates**.
- Collaborated with product and compliance teams to implement **data quality monitoring** and improve governance processes.
- Enhanced query performance and enabled **cross-team analytics** by integrating structured and unstructured datasets into Redshift.

UNT North Texas NOW! | Graduate Assistant

Jan 2025 - Present

- Analyzed **150+ institutional datasets**, uncovering key performance trends that informed academic retention strategies.
- Applied **Python (Pandas, NumPy)** for predictive analysis, supporting data-driven process improvements.

College of Applied & Collaborative Studies | Teaching Assistant

Aug 2024 - Present

- Supported **100+ students** in Python- and SQL-driven data analysis projects, improving project outcomes and participation by 15%.
- Guided teams on applying **data visualization tools** (Tableau, Power BI) to communicate insights effectively.

Upwork | Full Stack Software Engineer (Top-Rated)

Dec 2022 - Jan 2024

- Delivered **cloud-based applications** with Next.js, React, Django, and FastAPI on AWS, Firebase, and GCP.
- Achieved **100% Job Success Score** and **Top-Rated status** within two months, highlighting client satisfaction and reliability.
- Partnered with global clients to design **data-driven dashboards and reporting tools** that enhanced visibility into KPIs.

Deerwalk Groups | Developer Intern

Dec 2019 - Aug 2021

- Created **CI/CD pipelines**, reducing deployment times from hours to minutes and enabling zero-downtime updates.
- Supported development of **data visualization dashboards** and scalable backend systems for internal use.

TECHNICAL SKILLS

- Data Engineering:** AWS (S3, Redshift, Glue, Lambda, SES), ETL Pipelines, Web Scraping (Selenium, BeautifulSoup), Spark, Hadoop, HDFS
- Data & Analytics:** Python (Pandas, Numpy, Scikit-Learn, Matplotlib), SQL (MySQL, PostgreSQL, MongoDB), Tableau, Power BI
- AI & Applied ML:** Vertex AI (Gemini 2.5, OpenAI APIs), LangChain / LangGraph, RAG Pipelines, Vector Databases
- Cloud & DevOps:** Google Cloud Platform (Cloud Run, Storage Bucket, Vertex AI), Docker, Git, CI/CD, Vercel
- Software Development:** JavaScript (React, Next.js, Node.js, NestJS), FastAPI, PHP, Java, Flutter, KPI Dashboards

PROJECTS | [LINK](#)

LawgicBot – AI-Powered Legal Education Platform | [LINK](#)

- Built full-stack AI chatbot platform with FastAPI, LangChain, and Next.js; deployed on Google Cloud with zero-data-retention architecture for privacy.
- Enabled real-time legal document analysis and automated PDF generation, improving accessibility to legal resources for non-technical users.

Basketball Analytics – Tournament Prediction System | [LINK](#)

- Developed ML models (Random Forest, Logistic Regression, SVM) on 2,000+ NCAA team records, leveraging Pandas, Scikit-Learn, and SQL for feature engineering.
- Achieved 95% AUC for qualification predictions and 90% accuracy in upset detection, enabling actionable insights for scouting and performance analysis.

Big Data Health Risk Prediction | [LINK](#)

- Designed real-time heart attack risk predictor using Random Forest and SVM; deployed with FastAPI + Next.js + Vercel.
- Processed user health metrics (age, cholesterol, stress, exercise, BMI) to deliver instant AI-driven risk predictions, supporting preventive healthcare decisions.

EDUCATION

University of North Texas | Masters, Advanced Data Analytics (GPA: 4.0)

Jan 2024 - Present