

TAYOUTH MALLA

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

SUMMARY

Data Engineer and Full-Stack Developer with **5+ years of experience** in building **scalable ETL pipelines, real-time analytics,** and cloud-based **AI solutions**. Skilled in **Python, SQL, AWS, GCP, and Spark**, with hands-on projects in **predictive analytics** and **RAG-powered AI**. Pursuing a **Master's in Advanced Data Analytics**, passionate about turning complex data into actionable insights.

WORK EXPERIENCE

Verity Advisors

Feb 2025 - May 2025

Data Engineering Intern

- Engineered and optimized **ETL pipelines** using AWS Glue, Lambda, and Redshift to process **1M+ rows weekly**, reducing reporting latency by **40%**.
- Automated **real-time data ingestion** with APIs, Selenium, and BeautifulSoup; delivered **error-free daily reports** via AWS SES.
- Developed **compliance monitoring systems** with automated alerts, strengthening governance and risk controls.
- Improved query performance by integrating **structured and unstructured datasets** into Redshift, enhancing cross-team analytics.

UNT North Texas NOW!

Jan 2025 - Present

Graduate Assistant

- Analyzed **150+ program datasets** to uncover performance trends; recommended strategies that improved student retention.
- Applied **Python (Pandas, NumPy)** and SQL to provide **data-driven insights** for institutional decision-making.

College of Applied & Collaborative Studies

Aug 2024 - Present

Teaching Assistant

- Guided **100+ students** in data analysis projects using Python and SQL; improved participation and project outcomes by **15%**.

Upwork

Dec 2022 - Jan 2024

Full Stack Software Engineer (Top-Rated)

- 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 through consistent client satisfaction.

Deerwalk Groups Web

Dec 2019 - Aug 2021

Developer Intern

- Built **CI/CD pipelines**, reducing deployment times from hours to minutes, enabling **zero downtime releases**.
- Supported full-stack development projects integrating **data visualization** and scalable backend systems.

TECHNICAL SKILLS

- Data Engineering:** AWS (S3, Redshift, Glue, Lambda, SES), Spark, Hadoop, ETL Pipelines, GCP (Cloud Run, BigQuery, Vertex AI), Web Scraping (Selenium, BeautifulSoup)
- Data & Analytics:** Python (Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn), SQL (PostgreSQL, MySQL, MongoDB), R, Tableau, PowerBI
- Machine Learning & AI:** Supervised/Unsupervised ML, LangChain / LangGraph, RAG Pipelines, OpenAI APIs
- Cloud & DevOps:** Docker, Git, CI/CD (GitHub Actions, Jenkins), Vercel/Streamlit, Firebase
- Software Development:** JavaScript (React, Next.js, Node.js, NestJS), FastAPI, PHP, Java, Flutter, Django

PROJECTS | [LINK](#)

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

- Built a full-stack **AI chatbot** with **RAG based (FastAPI, LangChain)** and **Next.js frontend** for legal education and automated PDF generation using **AI agents**.
- Enabled real-time **legal document analysis** and **automated PDF generation** on **Google Cloud (Cloud Run, Vector DB)** with a **zero-data retention** architecture, delivering secure, accessible, and real-time legal resources for 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](#)

- Built a real-time **heart attack risk prediction system** using ML models (Random Forest, Logistic Regression, SVM) with deployment via **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

Jan 2024 - Present

Masters, Advanced Data Analytics

Denton, TX

- GPA:** 4.0