

Venkata Goutham Pachava

+1 (940) 365-6469 | Gouthampachava7@gmail.com | [GitHub](#) | [LinkedIn](#) | [Portfolio](#)

PROFESSIONAL EXPERIENCE

University of North Texas – Research Assistant

March 2024 – December 2024

- Automated personalized course ads with an **AI agent** across 4 platforms, optimizing posting times and content for student demographics, which led to increased enrollment growth.
- Built **Power BI dashboards (DAX)** to analyze **ROI** of online course marketing campaigns, improving **ROAS** by 25%.
- Streamlined clickstream reporting with **MS Excel macros/VBA & Power Query**, cutting data processing time by **48%** and eliminating manual errors.

Tata Consultancy Services – Data Analyst | Hyderabad, India

January 2021 – September 2023

Data Analyst (2022 – 2023)

- Designed & optimized **BigQuery SQL pipelines** for 1.3x faster data retrieval and transformation.
- Conducted **A/B testing** on 50+ campaigns, cutting bounce rate by **46%** and increasing engagement by **12%**, learning hands-on about statistical significance and experiment design.
- Collaborated on a **Random Forest churn prediction model**, improving retention by **25%** and driving **\$2.6M projected annual revenue**.
- Automated data quality checks via **Alteryx + spaCy ETL workflows**, saving **92% of manual effort**.
- Delivered interactive **Power BI reports**, reducing stock-to-sales ratio by **35%** and improving resource allocation by **45%**.

Junior Data Analyst (2021)

- Automated manual Excel-based reporting workflows using **Python & SQL**, reducing reporting time by **30%** and enabling real-time decision-making.
- Learned **data wrangling, data cleaning, and visualization**, building early **Power BI dashboards** to track monthly supply chain KPIs and uncover process bottlenecks.
- Assisted in **exploratory data analysis (EDA)** for internal projects, preparing datasets and visual summaries that supported senior analysts in hypothesis testing.

PROJECTS

OPTIMIZING INVESTMENTS ON AGRICULTURAL LAND – Personal Project - [GitHub](#) April 2025

- Conducted **exploratory data analysis (EDA)** on USDA datasets to uncover patterns in crop yields, soil characteristics, land value.
- Applied **statistical methods and regression models** to identify key factors influencing profitability.
- Designed **interactive Tableau dashboard** to visualize county-level yield trends, profitability forecasts.
- Delivered actionable **investment recommendations** for farmland acquisition, improving decision-making for crop and site selection. December

PETERBILT COST OPTIMIZATION ANALYTICS – UNT Hackathon-[GitHub](#) December 2024

- Secured **top 5** in the hackathon. Developed a highly efficient **Machine Learning** ensemble models.
- Achieved **88% accuracy** using **ensemble ML models (Random Forest, XGBoost, Gradient Boosting)**.
- Built **Tableau dashboards** to highlight cost trade-offs between high vs. low-cost configurations.

RESEARCH COPILOT – AI RESEARCH ASSISTANT – HackUNT 2024- [GitHub](#) October 2024

- Developed a **Generative AI**-driven research assistant leveraging **GPT-4, LangChain,** and **FAISS** for context-aware summarization and Q&A.
- Reduced research paper review time by **70%**, enabling professors and students to extract key insights and interact with papers in under 5 minutes.

EDUCATION

University of North Texas, **MS Advanced Data Analytics** August 2023 – May 2025

Relevant Coursework: Statistics, Deep Learning, Recurrent Neural Network, Agile for Analytics, Data Visualization

Sathyabama University, **Bachelor of Engineering in Computer Science** July 2017 – May 2021

SKILLS

Technical: C/C++, Python, Java, HTML, CSS, JavaScript, Go, R, Spark, Hadoop, Hive, Data Structures & Algorithms, Linux, MLflow, Tableau, Power BI, Git, GitLab.

AI/ML: Math, Algebra, Causal Inference, A/B Testing, Feature Engineering, Data Modeling, Clustering, PCA, EMR, CNN, RNN, LSTM, GCP, AWS, Sage Maker, NLP, LLM, TensorFlow, AI Agent, Classification, Regression, LangChain.

DB/Ops: Database Design, SQL, PL/SQL, Teradata, MySQL, MongoDB, BigQuery, Oracle, S3, Redshift, Vector DB, Flask, Django, CI/CD, Streamlit, Docker, CUDA, GPU, JIRA, ETL, Data Lake, Kafka, Data Warehouse.