

THANUSRI AENUGULA

+1 (940) 364-1654 | thanusriaenugula@tamu.edu | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Education

Texas A&M University, College Station

Master of Science in Data Science

Courses: Data Mining, Statistics, Mathematics for ML, Applied Analytics, Machine Learning, Deep Learning

Aug 2025 - May 2027

College Station, TX

Sri Sivasubramaniya Nadar College of Engineering

Bachelor of Engineering in Electrical and Electronics Engineering

Courses: Data Structures, Probability and Statistics, Database Management, Data Modeling, Artificial Intelligence

Aug 2021 - May 2025

Chennai, India

Experience

H-E-B | *Decision Science Intern*

May 2026 - Aug 2026

- Incoming Decision Science and Analytics Intern with the Customer Experience team, leveraging data analytics to enhance customer engagement and overall experience.

Texas A&M University | *Research Assistant, Advanced Vision and Learning Lab*

Dec 2025 - May 2026

- Establishing centralized data pipelines to integrate **3+ modalities** and analyze data for automated phenotyping across multiple tasks.
- Supporting analysis of **5+ plant growth traits** using AI and statistical modeling techniques enabling automated phenotyping.
- Implemented model evaluation and cross-validation frameworks to ensure trait prediction performance across diverse plant conditions.
- Collaborated with AgriLife plant research teams to translate phenotyping insights into actionable research decisions and experiments.

Daimler India Commercial Vehicles | *Data Science Intern*

Jan 2025 - Jun 2025

- Collaborated with the Data-Driven Engineering team to develop machine learning tools for the **cross-functional team**.
- Streamlined **5+** vehicle testing workflows by integrating predictive analytics into existing testing pipelines.
- Built a **statistical twin** for Dynamic Brake Testing systems using historical brake from **100+** vehicles enriching predictive analysis.
- Delivered a production-ready solution that enabled testing team to pre-screen vehicle performance, reducing testing time by **80%**.
- Investigated telematics and Dealer Management Service data to support driving pattern analysis and fleet optimization.
- Developed **ETL pipelines** in **Azure Databricks** for real-time monitoring and analytics.

Defence Research and Development Organisation (DRDO) | *Data Science Intern*

Sep 2024 - Nov 2024

- Integrated hybrid models to predict battery Remaining Useful Life and State of Charge, achieving **R² of 0.97** and enhanced accuracy.
- Examined performance data from **38+** lithium-ion batteries, reducing failure prediction time and identifying early degradation.
- Architected monitoring pipelines to track battery health helping the team plan maintenance and manage assets more effectively.
- Developed Power BI visualization dashboards to monitor degradation trends and support decision-making.

Projects

Multi-Agent SQL Query Execution System [[GitHub](#)]

- Designed a LangGraph-based **multi-agent** system automating SQL analytics and reporting, reducing manual effort by **60%**.
- Built a SQL reasoning pipeline with enforced grounding and summarization constraints, ensuring **100%** query execution validity.
- Deployed on **AWS EC2**, enabling real-time multi-agent inference < **2s** downloadable PDF report generation.
- Demonstrated SQL automation, producing reports from natural language queries and validating results across complex database.

Predictive Maintenance of Wind Turbine Systems [[GitHub](#)]

- Designed hybrid **CNN-LSTM-GRU** architectures for Remaining Useful Life (RUL) prediction of wind turbine systems.
- Attained an **R² of 0.91**, enhanced the reliability forecasting and predictive maintenance planning.
- Validated hybrid model, accurately predicting Remaining Useful Life and capturing turbine performance patterns.
- Crafted an interactive **Power BI** dashboard to monitor **10+ KPIs** and visualize system performance trends.

Corrective RAG System for AI Knowledge Retrieval [[Github](#)]

- Designed a system using **LangGraph** to validate retrieval context and correct queries, reducing unsupported LLM responses.
- Built an end-to-end RAG pipeline with web scraping, semantic chunking, and **FAISS** vector search for low-latency, grounded retrieval.
- Deployed a Streamlit application on **AWS EC2**, enabling real-time inference with **low response latency**.

Technical Skills

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

Frameworks & Libraries: PyTorch, TensorFlow, Scikit-learn, Hugging Face Transformers, LangChain, LlamaIndex, LangGraph, NumPy, Pandas, PySpark, Matplotlib, Seaborn

AI/ML Architectures: Supervised Learning, Unsupervised Learning, Predictive Modeling, Time Series Analysis, Anomaly Detection, Clustering, Classification, Regression, Ensemble Methods (XGBoost, Random Forest), NLP, Computer Vision

Deep Learning: CNN, RNN, LSTM, GRU, Autoencoders, Transformers, Large Language Models

Cloud & Big Data Tools: AWS (SageMaker, Bedrock, S3, EC2, Lambda), Azure Databricks, Apache Spark, Git, MySQL

Visualization Tools: Excel, Tableau, Power BI

Certification

The University of Texas at Austin [[Certificate](#)]

Post Graduate Program in Data Science and Business Analytics

Aug 2023 - Aug 2024

CGPA: 4/4