

EDUCATION

The University of Texas at Dallas	Jan 2025 - Dec 2026
Master of Business Analytics and Artificial Intelligence (STEM)	GPA 3.7
Scholarship: Analytics Excellence scholarship	
Vellore Institute of Technology	May 2019 - July 2023
Bachelor of Computer Science and Engineering	GPA 3.14

SKILLS

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

Data Analysis & Machine Learning: Data Preprocessing, Feature Engineering, ML Algorithms, Data Visualization, ETL, BI, Predictive & Prescriptive Modeling, Advanced Statistics, Generative AI, Prompt Engineering, Model Evaluation & Validation.

Data Science: Statistical Modeling, Data Mining, Predictive Modeling, Experimentation, Model Automation, Anomaly Detection

AI & ML Concepts: Responsible AI, Model Deployment Concepts, AI Security & Governance, Inference & Fine-Tuning Concepts

Tools & Platforms: Power BI, Docker, Jupyter Notebook, Microsoft SQL Server, Git, SSIS, SSAS, SSRS, Flask, Oracle Cloud Infrastructure

Big Data & Cloud Platforms: Hadoop (HDFS, Yarn), Spark (PySpark, Spark SQL), Hive, IBM Watson Studio, Azure

Libraries & Frameworks: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, PyTorch. Foundational knowledge of AWS and cloud-based data warehousing concepts, n8n (Automation).

PROFESSIONAL EXPERIENCE

LTIMindtree, Noida, India	January 2024 – December 2024
Data Engineer	
<ul style="list-style-type: none">Developed and automated large-scale data pipelines (1,100+ daily runs) using SQL and Azure, enabling efficient processing of high-volume datasets and supporting analytical modeling and data-driven decision-making.Designed and implemented data validation frameworks and anomaly detection systems to identify system errors, improving data quality and reliability for downstream statistical models.Translated complex data into actionable insights and presented findings to business stakeholders, influencing decision-making and improving operational efficiency.	
LTIMindtree, Mumbai, India	April 2023 – July 2023
Data Analyst Intern	
<ul style="list-style-type: none">Reduced operational inefficiencies and prevented revenue loss by designing Fact and Dimension tables with SSIS and implementing SCDs, ensuring high-quality, reliable data for business reporting and decision-making.Partnered with business stakeholders to develop interactive Power BI dashboards using SSAS and DAX, delivering actionable insights that improved promotion strategies and supported data-driven decisions.Built and optimized reporting solutions by integrating SSMS, SSIS, and SSAS, ensuring data accuracy, consistency, and governance, while translating complex data into clear insights for non-technical stakeholders.Owned end-to-end data workflows, including data extraction, preprocessing, modeling support, and insight delivery for large scalable analytical systems.	

ACADEMIC PROJECT EXPERIENCE

Big Data Analysis: NYC Yellow Taxi & Weather Correlation	Aug 2025 - Nov 2025
Technologies used: Hadoop (HDFS, YARN), Spark (PySpark, Spark SQL), Hive, Python, SQL	
<ul style="list-style-type: none">Processed and analyzed over 1GB of raw NYC Yellow Taxi and weather datasets using a multi-node Hadoop cluster, applying data mining techniques to identify demand patterns during extreme weather events.Optimized big data workflows by configuring YARN site settings and utilizing Spark DataFrames to perform complex joins and correlation analysis across millions of records.Extracted and mapped weather station IDs to taxi zones to provide statistically grounded insights into how precipitation and temperature fluctuations impact urban mobility.	
End-to-End Retail Sales Analytics & Demand Modeling	
Technologies used: Python, Pandas, NumPy, Statistical Modeling, ML (OLS, Lasso, Ridge, Elastic Net)	Oct2025-Dec 2025
<ul style="list-style-type: none">Engineered and transformed 167K+ rows of retail scanner data, creating 54 high-impact features to build scalable, model-ready datasets for large-scale analytical modeling.Built and validated statistical and machine learning models (OLS, Lasso, Ridge) using data mining and cross-validation techniques, uncovering key sales drivers and enabling pricing and promotion strategies that mitigated a 30% sales decline.Owned end-to-end modeling workflows, from data preprocessing to model evaluation, and translated complex analytical insights into clear recommendations for stakeholders, driving data-informed business decisions.	

CERTIFICATIONS

- Oracle OCI Data Science Professional | October 2025
- Oracle Cloud Infrastructure 2025 Generative AI Professional | November 2025
- Salesforce Certified Agent force Specialist | November 2025
- Big Data Analytics | University of California

LEADERSHIP & MENTORSHIP

Student Mentor, University of Texas at Dallas	Nov 2025 – Present
---	--------------------