ROSHAN SIDDARTHA SIVAKUMAR

615-364-5138 | roshan.siddartha.sivakumar@vanderbilt.edu | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Data science graduate student with experience in data engineering, machine learning, and stakeholder-facing projects. Recently led a three-month industry collaboration with a global automotive client, developing tools and dashboards that identified opportunities to improve packaging efficiency by over 20%. Passionate about building scalable, insight-driven solutions that drive measurable results.

EDUCATION

Vanderbilt University – Data Science Institute

Master of Science in Data Science (GPA: 3.65/4)

Relevant Coursework: -Probability and Statistical Inference, Survey of Data Science Applications, Exploratory Data Analysis, Principles of Programming and Simulation, Data Science Rights Responsibility, Machine Learning, Data Management Systems, Data Science in Teamwork Practice

Vellore Institute of Technology

Bachelor of Technology (GPA: 8.47/10.0)

Major: Mechanical Engineering

Relevant Coursework: Problem solving and Programming, Problem Solving and Object-Oriented Programming, Statistics for Engineers, Linear Algebra and Differential Equations, Programming Data Structures and Algorithms using Python

SKILLS

Relevant Skills: Machine Learning (Model Training & Evaluation), Data Cleaning & Transformation, Feature Engineering, Data Analysis, Data Visualization, ETL Workflows, Model Testing & Validation, Statistical Modeling, Problem Solving, SQL Querying, Data Modeling & Normalization, Audit Logging, Stored Procedures & Transactions

Related Platforms: Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn), SQL (MySQL), Jupyter, GitHub, Tableau, AWS, Docker, Google Colab, Microsoft Office 365, PHP (basic)

PROJECTS

Packaging Optimization for Automotive Supply Chain (Confidential)

Team Lead | *Vanderbilt University x Nissan (Teamwork in Data Science)*

- Led a packaging optimization project in collaboration with Nissan, focusing on improving packaging efficiency across part families using structured part dimension and material usage data.
- Cleaned and engineered highly incomplete datasets (~99% null) using Python, including part grouping logic and imputation. •
- Developed interactive Tableau dashboards exposing inefficiencies in space usage and packaging materials.
- Delivered insights projected to enable significant cost savings through improved standardization and visualization.

Sales Playbook Optimization Dashboard

Team Lead | Vanderbilt University (Machine Learning Course)

- Constructed a predictive classification pipeline using ensemble models to forecast B2B deal outcomes with >98% accuracy.
- Merged and transformed CRM data into model-ready formats with engineered behavioral and demographic features.
- Deployed a Streamlit dashboard using Docker for real-time deal scoring and win-rate monitoring. •
- Enhanced pipeline robustness through iterative tuning and stratified sampling to address class imbalance.

Economic Freedom & Societal Outcomes Analysis

Team Lead | *Vanderbilt University (Exploratory Data Analysis)*

- Investigated global economic patterns using PCA, cluster analysis, and regression techniques on freedom, GDP, and wellbeing indicators.
- Segmented 160+ countries into four economic clusters using unsupervised learning to reveal structural development patterns.
- Visualized complex relationships between economic freedom, happiness, healthcare, and food affordability across topperforming nations.
- Derived policy-level insights showing that economic liberty alone does not guarantee improved societal outcomes. Nashville, TN

Voter Database Design & Normalization Project

Individual Contributor | Vanderbilt University (Database Management Systems) Jan 2025 – Apr 2025

- Structured and normalized a ~4GB MySQL database using functional dependencies and BCNF to support scalable voter • registration workflows.
- Coded stored procedures, triggers, transactions, and constraints to enforce data integrity and auditability.
- Linked the backend to a PHP-based interface enabling real-time voter lookup, registration, and analytics. •
- Generated SOL views and audit tables to maintain change history and promote system-wide transparency.

ADDITIONAL

- Language Skills: English, Tamil and Telugu
- Certifications: Certificate of Recognition Nissan & Vanderbilt Packaging Optimization Project (Apr 2025), Programming, Data Structures and Algorithms Using Python - NPTEL India, AI Analyst - IBM
- Interests: listening to contemporary pop music, playing story-driven and open-world games, exploring the latest advancements in AI

Nashville, TN

Jan 2025 – Apr 2025

Jan 2025 – Apr 2025

Nashville, TN

Nashville, TN

Oct 2024 – Dec 2024

Vellore, India

August 2024 - May 2026

Nashville, TN

July 2019 – June 2023