

SUSHEEL SRIKANTH

+1 (615) 414-3621 | susheel.srikanth@vanderbilt.edu | www.linkedin.com/in/susheelsrikanth | github.com/Susheel1101

EDUCATION

Vanderbilt University – Data Science Institute

Nashville, TN

Master of Science in Data Science (GPA: 3.962/4.00)

August 2023 – May 2025

- Relevant Coursework: Deep Learning, Generative AI Models, Machine Learning, NLP in Asset Management, Data Science Algorithms, Big Data Scaling, Statistics, Exploratory Data Analysis
- Teaching Assistantships: Business Analytics, Intro to Data Science, Principles of Programming, Database Management Systems

Vellore Institute of Technology

Vellore, India

Bachelor of Technology in Mechanical Engineering (GPA: 3.55/4.00)

July 2019 – April 2023

- Relevant Coursework: Linear Algebra, Object-Oriented Programming, Total Quality Management, Consumer Behavior

EXPERIENCE

SymTrain (Capstone Project)

Nashville, TN

Data Scientist

August 2024 – December 2024

- Built an AI-powered data solution to automate training simulations, reducing onboarding time by 70% through process optimization and business analysis
- Designed computer vision solution to detect UI interactions across 5,000+ frames, enabling automated interface analysis and improving simulation relevance
- Developed bounding-box algorithms for cursor mapping, increasing simulation realism by ~40% (user feedback)
- Collaborated cross-functionally with product teams to define simulation requirements and deliver scalable solutions on AWS

Stardust Materials LLC

Vancouver, WA

Data Science Intern

May 2024 – August 2024

- Boosted secure product tracing model accuracy by 20% using photoluminescence simulations across specialized materials
- Extended quantum chemistry pipelines using PySCF, CP2K, and Quantum Espresso for 20+ molecular configurations
- Reduced simulation runtime by ~30% through GCP-based distributed deployment, enabling large-scale testing

Data Science Institute, Vanderbilt University

Nashville, TN

Research Assistant

February 2024 – April 2024

- Supported early-stage policy research on the effects of conservation efforts on federal transfers, election outcomes, and candidate quality in Brazil
- Helped build a comprehensive dataset on environmental policy and local political dynamics, using R and excel to clean, merge, and transform 50K+ records
- Ensured data accuracy and contextual relevance for downstream modeling and statistical analysis

PROJECTS

Stock Return Prediction Using Financial News (Alliance Bernstein)

January 2025 – April 2025

Tools: Python, BERT, LLAMA3, XGBoost, Random Forest

- Built a pipeline to forecast short-term stock returns (1-, 5-, 10-day) by combining news-based sentiment with market indicators
- Used LLAMA3 and FinBERT to extract sentiment signals from news; merged with historical pricing trends and volatility metrics
- Performed data analysis and trained tree-based models, achieving 63% accuracy, validated performance through Sharpe ratio and back testing

AI – Powered Job Matching Platform – Vanderbilt University

February 2025 – April 2025

Tools: Python, FAISS, AWS (EC2, S3), OpenAI, Flask, Langfuse

- Developed a job matching system using both deterministic profiling and semantic retrieval to enhance user-fit recommendations
- Built an embedding index with FAISS and deployed the system on AWS EC2 + S3, delivering ~6.7s latency per query
- Tested multiple orchestration strategies (CrewAI vs MultiAgent), improving match quality by 20% while lowering API cost

Sentiment Analysis of Amazon Product Reviews – Vanderbilt University

January 2025 – April 2025

Tools: PySpark, GCP, TF-IDF, Word2Vec, Logistic Regression, Random Forest, NLP

- Applied data mining and sentiment analysis techniques to extract insights from 18M+ product reviews, supporting category-specific product strategies
- Delivered business insights to support category-specific product strategies using sentiment analysis
- Performed data analysis and model tuning to support marketing optimization, improving customer understanding

SKILLS

Languages & Libraries: Python (Pandas, NumPy, Scikit-learn, PyTorch), SQL, R, PySpark, Spark

Machine Learning & Analytics: Regression, Classification, Time-Series, NLP, EDA, Model Deployment, Fraud Detection

Cloud & Big Data: AWS (Lambda, S3, EC2), GCP (VMs), FAISS, Docker, Hadoop

Tools & Visualization: Tableau, Streamlit, Langfuse, Git, JIRA, Excel