LAKSHMI SWATHI SREEDHAR

DATA SCIENTIST – AI/ML

Lansing MI, 48917 · (313) 529-2810 · swathisl@umich.edu

LinkedIn | GitHub | Blog | HackerRank | Portfolio

PROFILE SUMMARY

Data Scientist with 2+ years of experience transforming complex data into smart, scalable solutions using AI and machine learning. Developed LLM-powered Text2SQL systems, BERT-based deduplication models, and multimodal AI tools impacting over 10,000 users and improving decision accuracy by up to 35 percent. Expertise in coding with Python, SQL, PySpark, and leveraging no-code innovation to deliver fast, effective results. Hands-on experience with cloud platforms GCP and Azure, focusing on data quality and automation. Currently expanding skills through MIT x PRO's No Code AI program. Seeking growth into solution engineering and architecture roles.

PROFESSIONAL EXPERIENCE

- **CHAINSYS CORPORATION** **09/2023 – 05/2025** Grand Ledge, MI

  Data Scientist – AI/ML & Data Engineering

  - Spearheaded the development of LLM-based Text2SQL systems and RAG pipelines, enhancing natural language querying and SQL precision for over 10,000 enterprise users, improving query efficiency by 30%.
  - Integrated Agentic AI with Text2SQL workflows to enable autonomous query generation and execution, reducing query time by 20% in enterprise applications.
  - Designed and deployed advanced AI models including a BERT-based deduplication algorithm and multimodal AI for healthcare, improving data quality by 16% and diagnostic precision by 35%.
  - Built real-time analytics dashboards and collaborated with 25+ engineers to improve prompt design for large datasets, boosting decision-making efficiency by 20% across 15+ clients.
  - Automated ETL workflows for 100,000+ daily records and migrated 1TB+ datasets across projects, cutting manual effort by 40% while ensuring accuracy and reliability.
  - Enhanced LLM performance through advanced prompt engineering techniques, ensuring accurate, scalable outputs for enterprise-scale data applications.
  - Implemented enterprise-grade data warehousing pipelines and data quality governance with SQL rule validation and AI-driven profiling, improving trust and data compliance.
  - Championed A/B testing protocols for AI applications, increasing active users by over 1,000 in one quarter.
  - Delivered $5M+ in tailored AI solutions, led A/B testing protocols that boosted active users by 1,000+, and developed GAN pipelines to increase dataset diversity by 50%.
  - Designed innovative data modeling strategies, delivering $1M+ in annual financial gains for enterprise clients through optimized data structures and workflows.

- **PARAILLEL INC** **03/2023 – 08/2023** Detroit, MI

  Machine Learning Developer Intern

  - Built NLP pipelines for sentiment analysis and text classification, supporting 50,000+ users.
  - Crafted and deployed targeted recommendation systems, increasing student engagement and retention rates by 30% among a diverse user base of over 20,000 active learners.
  - Incorporated federated learning frameworks, ensuring privacy-compliant AI solutions for sensitive educational data across 5+ global clients.
  - Engineered adaptive learning models, that boosted content recommendation outcomes by 20%.
  - Systemized content tagging and classification system using NLP techniques, reducing manual effort for educators by 40% across 50,000+ lessons.

EDUCATION

MIT X PRO – MASSACHUSETTS INSTITUTE OF TECHNOLOGY **02/2025 – 06/2025** REMOTE

No-Code AI and Machine Learning: Building Data Science Solutions

UNIVERSITY OF MICHIGAN DEARBORN **09/2021-04/ 2023** Dearborn, MI

Masters in Artificial Intelligence; Minor in Machine Learning

- Non-Resident Graduate Scholarship Recipient

DAYANANDA SAGAR COLLEGE OF ENGINEERING **08/2016 – 08/ 2020** Bangalore, India

Bachelors in Aeronautical Engineering

SKILLS

Programming & Scripting: Python, SQL, PySpark, R, Scala
Machine Learning & AI Frameworks: Scikit-learn, TensorFlow, PyTorch, Keras, XGBoost, Reinforcement Learning, Recommender Systems, Model Evaluation, A/B Testing, Bayesian Statistics
Deep Learning & Neural Networks: CNNs, RNNs, Transformers, GANs, VAEs, BERT, LLMs
NLP & Language Systems: Hugging Face, LangChain, Prompt Engineering, RAG Pipelines, TF-IDF, Topic Modeling, Agentic Systems, Explainable AI (XAI)
Data Engineering & Governance: ETL, Data Integration, Data Quality, Metadata Management, Data Governance
Tools & Platforms: Jupyter, Pandas, NumPy, Matplotlib, GitHub, Tableau, RapidMiner, KNIME
Cloud & Deployment: Google Cloud Platform (GCP), Microsoft Azure

PROJECTS / OPEN-SOURCE

- ***GAN-Based Customer Data Generation*** | ***Link*** *Python, PyTorch, NumPy, Matplotlib, Pandas*
  Developed a GAN generating 5,000 realistic synthetic customer data samples, improving data diversity and realism over rule-based methods.
- ***Autonomous-Driving-TurtleBot*** | ***Link*** *MATLAB*
  Created an autonomous driving algorithm with ROS, image processing, and Lidar, enabling lane detection and obstacle avoidance.
- ***Health Monitoring Alert System*** | ***Link*** *Python, Pandas, NumPy, Scikit-learn, Matplotlib*
  Built a Random Forest classifier to predict health alerts with 90%+ accuracy and analyzed key features.
- ***AI-Powered UNSPSC Category Generator*** | ***Link*** *Python, OpenAI GPT-3.5 API, Pandas, Excel, CSV*
  Automated generation of 10,000+ UNSPSC category titles using GPT-3.5, ensuring scalability and performance.
- ***Semantic Segmentation for Autonomous Driving*** *Python, PyTorch, TensorFlow, OpenCV, Cityscapes Dataset, ROS (Robot Operating System)*
  Developed a semantic segmentation model for autonomous driving, achieving >85% IoU and real-time speeds for detecting roads, vehicles, pedestrians, and lanes.

CERTIFICATIONS

- Become a Data Scientist - LinkedIn
- Train and manage a machine learning model with Azure Machine Learning - Microsoft
- Run pipelines in Azure Machine Learning - Microsoft
- Machine Learning Specialization by Stanford, Andrew NG - Coursera
- Academy Accreditation – Generative AI Fundamentals - Databricks
- Creating Multi Task Models with Keras - Coursera
- No Code AI and Machine Learning: Building Data Science Solutions by MIT professional Education – Ongoing

PUBLICATIONS AND PRESENTATION

- ***Advancing Precision Medicine through Multimodal AI: Innovative Approaches to Diagnostics and Treatment***

Proposes a multimodal AI framework combining medical imaging, clinical text, and wearable data to improve diagnostics and personalized treatments while ensuring data privacy with federated learning.

- ***Speak the Language of AI: Mastering Prompt Engineering for LLMs***

Explores advanced prompt engineering techniques for LLMs, introducing automated tuning frameworks to improve performance across business, education, and healthcare.