LAKSHMI SWATHI SREEDHAR

Lansing MI, 48917 · 313-529-2810 · swathisl@umich.edu
LinkedIn | GitHub | Blog | HackerRank | Portfolio

PROFILE SUMMARY

AI Data Scientist with a strategic product mindset and a builder's bias for action, transforming raw data into intelligent, scalable, user-centric systems. Over 2+ years of hands-on experience developing LLM-powered Text2SQL platforms, BERT-based deduplication engines, and multimodal AI tools, driving up to a 35% improvement in decision accuracy and positively impacting 10,000+ users across diverse industries. Proficient in Python, SQL, PySpark, and cloud platforms including GCP and Azure, I combine deep technical expertise with agile execution to deliver fast, scalable solutions leveraging both code and no-code tools that truly move the needle. A lifelong learner committed to continuous upskilling, I've completed MIT x PRO's No-Code AI program and am passionate about bridging data science with product leadership, ready to own AI strategy, lead cross-functional teams, and architect scalable, future-forward solutions.

PROFESSIONAL EXPERIENCE

- **CHAINSYS CORPORATION** **09/2023 – 05/2025** Grand Ledge, MI
  Data Scientist – AI/ML & Data Engineering

- Founded and led development of a multi-agent AI framework enabling autonomous capabilities across multiple products; drove the project from concept to shipment within months while leading cross-functional teams.
- Served as Product Team Lead, coordinating AI engineers, SQL developers, and QA teams; facilitated agile workflows and sprint planning to align features with business goals and accelerate delivery.
- Led mentorship and knowledge-sharing initiatives to maintain high standards and foster team growth.
- Spearheaded LLM-powered Text2SQL and Retrieval-Augmented Generation (RAG) pipelines, improving natural language querying and SQL precision for 10,000+ enterprise users and boosting query efficiency by 30%.
- Integrated Agentic AI in Text2SQL workflows to enable autonomous query generation, reducing response time by 20%.
- Automated ETL workflows processing 100,000+ daily ERP/business records and managed 1TB+ data migrations, cutting manual effort by 40% while maintaining data accuracy and integrity.
- Designed scalable data warehousing and governance frameworks using SQL rule validation and AI-driven profiling to enforce quality and compliance across ERP/CRM systems.
- Collaborated with ERP integration and cross-functional teams to optimize data flows and system interoperability within the ChainSys Smart Data Platform™.
- Worked extensively with PostgreSQL, Oracle, and Databricks SQL for data management and querying across projects.
- Applied Kafka basics for real-time data streaming integration and used Tableau for building insightful visualization dashboards with strong UI/UX design principles focused on user experience for AI tools.
- Developed intuitive UI/UX for AI framework components and dashboards, ensuring seamless interaction and usability for end users and stakeholders.
- Applied advanced prompt engineering to fine-tune LLM outputs for scalable, accurate enterprise query generation.
- Led A/B testing initiatives for AI tools, boosting active user engagement by over 1,000 in one quarter.
- Delivered customized AI and data engineering solutions valued at $5M+, driving operational excellence and financial impact for enterprise customers.

- **PARAILLEL INC** **03/2023 – 08/2023** Detroit, MI
  Machine Learning Developer Intern

- Acted as Scrum-inspired team lead within the ML internship program, facilitating agile workflows, sprint planning, and daily stand-ups to ensure timely delivery and effective team collaboration.
- Fostered a culture of continuous learning and knowledge sharing, enabling the team to overcome challenges and maintain high-quality standards.
- Built robust NLP pipelines for sentiment analysis and text classification, supporting a user base of over 50,000 learners.
- Designed and deployed targeted recommendation systems that increased student engagement and retention by 30% among 20,000+ active users.
- Incorporated federated learning frameworks to deliver privacy-compliant AI solutions for sensitive educational data across 5+ global clients.
- Engineered adaptive learning models that improved content recommendation outcomes by 20% and developed a systematic content tagging and classification system, reducing manual educator workload by 40% across 50,000+ lessons.

- **APTECH INC** **12/2020 – 04/2021** **BANGALORE, INDIA**
  AI/ML Engineer Trainee

- Completed 408 hours of intensive instructor-led training covering AI/ML fundamentals, software development, and advanced product engineering concepts.
- Built Python-based machine learning pipelines for classification, clustering, regression, and time series forecasting using Scikit-learn.
- Developed RESTful APIs and GUI applications with Flask and PyQt to integrate ML models and visualize results.
- Performed data preprocessing, exploratory data analysis (EDA), and dimensionality reduction using Pandas, NumPy, and Matplotlib.
- Applied core mathematical and statistical concepts including linear algebra, probability, and multivariable calculus to support model development and optimization.
- Gained hands-on experience in agile workflows, version control with Git, and real-time debugging within a simulated production environment.

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 (PostgreSQL, Oracle, MySQL, Databricks), PySpark, R, Scala, Bash
Machine Learning & AI: Scikit-learn, XGBoost, LightGBM, TensorFlow, PyTorch, Keras, Reinforcement Learning, Recommender Systems, A/B Testing, Bayesian Methods
Deep Learning & LLMs: Transformers, CNNs, RNNs, GANs, VAEs, BERT, Fine-tuning, Text2SQL, Multimodal Models, Agentic AI
NLP & Language Technologies: Hugging Face, LangChain, Prompt Engineering, RAG Pipelines, Topic Modeling, TF-IDF, Explainable AI (XAI)
Data Engineering & Governance: ETL, Apache Airflow, Kafka, Data Integration & Quality, Metadata Management, Azure Data Factory, SAP & Oracle ERP Pipelines
MLOps & Deployment: Streamlit, Flask, Docker, MLflow, API Integration, CI/CD, Model Monitoring, Experiment Tracking
Visualization & UI/UX: Tableau, PyQt, Dash, Streamlit, Jupyter, Matplotlib, Pandas, NumPy**,** User-centered design for AI systems, dashboards, low-code workflows
Cloud & Platforms: Google Cloud (Vertex AI, BigQuery), Microsoft Azure (Azure OpenAI, Synapse), SQLite, Ikigai, Google Teachable Machine
Development Methodologies & Tools: Agile, Scrum, Jira, Git, GitHub, Confluence, Test-Driven Development, Real-time Debugging, Cross-functional Collaboration

---

PROJECTS / OPEN-SOURCE

- ***AI-Powered Text2SQL and RAG System*** *Python, OpenAI GPT-4, LangChain, SQL, Streamlit, PostgreSQL, Oracle, Databricks, RAG, Agentic Workflows*
- Built a production-grade LLM-based Text2SQL system with RAG and agentic workflows, boosting SQL precision by 30% and enabling natural language queries across enterprise databases with prompt tuning and feedback loops.
- ***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 analysed 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 –MIT

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.