

SOWMYA SIDDANKI

 sowmyasiddanki@gmail.com  +1 (980)-315-1299  Boca Raton, Florida

 <https://www.linkedin.com/in/siddankisowmya/>

I am a Data Science graduate student at Florida Atlantic University with hands-on experience building end-to-end machine learning and data-driven applications. I work with Python, SQL, Power BI, and cloud analytics to turn complex data into clear, actionable insights. I am seeking a Data Science or Machine Learning role where I can contribute to real-world problems.

EDUCATION

Master of Science in Data Science and Analytics | GPA: 3.9/4.00

Florida Atlantic University

2024 – 2026 | Boca Raton, USA

- Completed courses in Data Mining and Machine Learning, Deep Learning, Natural Language Processing, and Computational Foundations of AI.

Bachelor of Technology in Computer Science and Engineering (Data Science) |

GPA: 7.83/10.00

2020 – 2024 | Hyderabad, India

Institute of Aeronautical Engineering

- Established a strong foundation in programming, data structures, and data science fundamentals, demonstrated by finalist placement in a nationwide hackathon.

PROJECTS AND WORK EXPERIENCE

Customer Segmentation using ML models for E-commerce websites

09/2024 – 12/2024

- Designed an **end-to-end cloud-based customer segmentation pipeline** using Microsoft Azure, with data ingestion and orchestration handled through **Azure Data Factory** and storage in **Azure SQL Database**.
- Built and trained **K-Means clustering models** using **Azure Machine Learning**, optimizing the number of clusters with **Silhouette Score and Elbow Method** to improve segment clarity and interpretability.
- Visualized customer segments and business insights using **Azure analytics** workflows, while enabling **scalable and automated model execution** without relying on local environments.

Sentiment-Driven Recommendation System for E-Commerce

02/2024 – 06/2024

- Built a scalable recommendation system using collaborative filtering and NLP on **50,000+ customer reviews**, improving recommendation relevance by **~30%** through techniques like **TF-IDF and cosine similarity**.
- Designed solutions for **cold-start users and sparse data**, evaluated recommendation quality using similarity and relevance metrics, and improved product coverage and personalization.
- Deployed an interactive **Streamlit** application for real-time, multi-user recommendations; optimized pipelines to reduce latency by **25%** and increased product discoverability and user engagement by **~20%**.

Deep Learning-Based Retrieval Augmented Generation (RAG) System

01/2025 – 04/2025

- Built a **deep learning-based RAG system** using **Python, transformer-based embedding models, and a GPT-based large language model** to answer questions from large unstructured document collections.
- Implemented **document chunking, embedding generation, and vector similarity search** using **PostgreSQL with vector extensions** to retrieve the most relevant text passages for each query.
- Integrated retrieved context with the **GPT-based model** to generate accurate, context-aware responses, reduced hallucinations, and evaluated performance using relevance, coherence, and consistency metrics.

SKILLS

Programming Languages: Python, SQL, JavaScript, Java, shell script

Web & Application Technologies: React.js, HTML, CSS, Streamlit

Databases: Microsoft SQL Server, Oracle, MySQL, PostgreSQL, MongoDB

Data Analytics & Visualization Tools: Pandas, NumPy, Scikit-Learn, Power BI, Tableau

Machine Learning, Deep Learning & LLM Technologies: TensorFlow, Keras, Transformer Models, Embedded Models

Cloud Platforms & Software Development: Microsoft Azure, Azure Data Factory, Azure Machine Learning, Azure Storage accounts, Git, GitHub, Azure SQL, Azure Databricks, Azure EventHubs, Agile Methodology (Azure DevOps), ServiceNow, Jira, Bitbucket, Atlassian,