

# NIVETHINI SETHILSELVAN

+1 (617) 206 0530 | [senhilselvan.n@northeastern.edu](mailto:senhilselvan.n@northeastern.edu) | [linkedin-nivethini-senthilselvan](#) | [GitHub](#) | [Tableau](#) | [Portfolio](#)

## EDUCATION

**Northeastern University - Master of Professional Studies in Applied Machine Intelligence (GPA – 3.9/4.0)**

**Boston, MA**

**Key courses:** Data Mining, AI Communication and Visualization, ML Operations, Business Intelligence.

**Dec 2025**

**Anna University - Bachelor of Technology in Information Technology (CGPA – 9.1/10)**

**Chennai, IN**

**Key courses:** Probability and Statistics, DBMS, Python, Java, OOPS, Data Structures, Supply Chain Management.

**May 2022**

## PROFESSIONAL EXPERIENCE

**Mutlicoreware/Uhnder Pvt Ltd**

**Chennai, IN**

**Software Engineer**

**Jun 2022 – Jul 2024**

- Developed a Performance Analysis Dashboard for CPU, RAM, DSP, and ACP usage, integrating **SQL** for data retrieval and processing real-time live data, leveraging **Matplotlib** and **Seaborn** for dynamic visualization.
- Engineered a Peer's KPI Metrics Dashboard to track bug metrics, test case execution, and automation coverage. Optimized data extraction with **SQL**, automated preprocessing with **Python**, and enabled real-time visualization in **Grafana**.
- Automated Radar performance data collection and integration with **SQL**, developed a **Flask-based dashboard** for real-time visualization, identifying 25% more undetected bugs across releases.

**Mutlicoreware/Uhnder Pvt Ltd**

**Chennai, IN**

**Intern - Software**

**Sept 2021 – May 2022**

- Developed a **Python Auto-Mail Trigger Script** to identify MISRA-C++ violations in Git commits.
- Developed a Hardware Inventory Dashboard using **Python/Flask** to display radar details and its current operational state.

## PROJECTS

**CPS-AI Assistant (Northeastern University) | RAG | Vector DB | OLLAMA | Groq | Streamlit (Github)**

**Feb 2025**

- RAG-Powered Chatbot – Built an AI chatbot using **Retrieval-Augmented Generation (RAG)** for precise, context-aware student queries.
- Efficient Vector Search – Implemented **Ollama embeddings with Supabase** vector DB for fast, accurate semantic retrieval.
- Scalable AI UI – Developed a **Streamlit-based interface** with real-time responses via **Groq's LLAMA 3.3 70B** model.

**Ask Your PA - AI Personal Assistant | GPT-3.5 | OpenAI Vector Embedding | Pinecone Vector DB (Github)**

**Apr 2025**

- Built a full-stack AI personal assistant using **FastAPI**, **Streamlit**, **GPT-3.5**, and **Pinecone** to deliver context-aware responses based on user-uploaded documents.
- Implemented secure user authentication, document parsing (.pdf, .docx, .txt), and OpenAI powered embedding pipeline for personalized vector storage and retrieval.
- Integrated a Streamlit-based chat UI with GPT to enable real-time querying of user-specific knowledge via **top-k cosine similarity** search from Pinecone.

**VibeTune - Mood Based Music Recommender | AWS MSK Kafka | EC2 | Glue | Redshift | Athena | TF-IDF (Github)**

**Mar 2025**

- Engineered real-time ingestion pipelines using **AWS MSK Kafka** and **S3** to stream and stage multi-source Spotify metadata (tracks, artists, features).
- Built distributed **ETL workflows in AWS Glue (PySpark)** to normalize, transform, and load data into a structured S3 data lake, enabling querying via **Athena** and **Redshift**.
- Implemented a content-based recommender using **TF-IDF** vectorization and **cosine similarity** on track-tag and lyric embeddings, deployed with a responsive Streamlit UI.

**Predictive Analytics for High-Value Customer Churn in the Telecom Sector | EDA & Machine Learning (Github)**

**Jul 2024**

- Developed a machine learning pipeline leveraging **logistic regression and decision tree classifiers** to predict customer churn for high-value telecom subscribers using monthly usage data.
- Feature Engineering & Dimensionality Reduction:** Applied **PCA** and advanced feature extraction techniques to optimize model performance and reduce dimensional complexity, handling class imbalance using **SMOTE**.
- Identified key churn indicators like call volume, data usage, and recharge frequency for retention strategies.

## TECHNICAL SKILLS

**Programming Languages:** Python (Pandas, NumPy, Matplotlib, Scikit-learn, Seaborn, TensorFlow, PyTorch), Java, SQL.

**Data Science and ML:** Regression, Classification, Decision Trees, SVM, Clustering, Neural Networks, NLP, LLM.

**Data Analytical Tools:** Tableau, Power BI, MS Excel (VLOOKUP, Pivot Tables, VBA), Jupyter Notebooks, Google Analytics.

**Data Warehouse Tools:** MySQL, MSSQL Server, GCP, AWS (S3, EC2), Snowflake, ETL, Airflow, Kubernetes, Docker.

**Other Tools and OS:** Jira, Confluence, Git, Agile, Kafka, MS PowerPoint, MS Word, Microsoft Office Suite, Windows, Linux.

**Soft Skills:** Strong Written and Verbal Communication, Easy Collaboration and Teamwork, Problem Solving, Analytical Skills.