Priyadharshan Sengutuvan

sengutuvan.p@northeastern.edu | (857) 397-8847 | linkedin.com/in/priyadharshan-sengutuvan

EDUCATION

Northeastern University, Khoury College of Computer Sciences — Boston, MA
Master of Science in Data Science — May 2026
Relevant Coursework: Machine Learning, NLP, Algorithms, MLOps

Sri Krishna College of Engineering and Technology — Coimbatore, India
Bachelor of Technology in Information Technology — May 2023

SKILLS

Programming: Python (NumPy, Pandas, Scikit-learn, PyTorch, TensorFlow), SQL (PostgreSQL, MySQL), R, C#
Machine Learning: Supervised/Unsupervised Learning, Deep Learning (CNNs, Transformers, BERT), NLP, XGBoost, Feature Engineering, Hyperparameter Tuning, Model Evaluation, A/B Testing
Data Engineering: Apache Spark, GCP BigQuery, Apache Airflow, Docker, ETL Pipelines, Neo4j, Git
Tools & Frameworks: FastAPI, LangChain, Tableau, Azure DevOps, CI/CD

EXPERIENCE

AriesView — Boston, MA (Remote)
AI/ML Engineer Intern — Sept 2025 – Dec 2025
- Built financial modeling platform with automated dashboards enabling investors to analyze property performance through real-time NOI, cap rate, and DSCR visualizations, deployed for 1K+ expected users.
- Created backend pipelines automating multi-scenario projections with sensitivity analysis to accelerate investment workflows for analysts.
- Enhanced OCR extraction system for financial documents by adding validation rules, boosting accuracy while eliminating manual data entry.
- Optimized RAG chatbot using Neo4j graph database for relationship-aware retrieval, experimenting with embeddings and hybrid search strategies.

Psiog Digital Private Limited — Chennai, India
Software Engineer Intern — Feb 2023 – Dec 2023
- Developed automated dashboards in React/C# displaying KPIs, replacing manual Excel workflows to enable data-driven decisions.
- Refactored SQL Server through query rewriting and indexing, significantly reducing load times for transaction analytics.
- Designed REST APIs connecting backend services, streamlining access across siloed systems.

PROJECTS

RAG-Based Operations Assistant for E-Commerce (GitHub) — *Python, LangChain, GCP, Airflow, Docker*
- Built RAG chatbot using LangChain and GPT-4 to automate seller support, leveraging semantic search with OpenAI embeddings.
- Architected Airflow ETL pipelines ingesting multi-source data into BigQuery with incremental loading for real-time dashboards.
- Improved retrieval precision from 0.68 to 0.87 through systematic experiments with embeddings, chunking, and prompt engineering.

F1 Race Position Predictor (GitHub) — *Python, FastAPI, Parquet, MLflow, Docker*
- Processed 26K+ F1 races across 14 datasets (1950-2024) achieving unified analytics schema by building ETL pipeline merging drivers, constructors, circuits, and qualifying data into Parquet format.
- Generated 47 time-series features reducing pipeline execution time by implementing vectorized rolling window aggregations, circuit-specific calculations, and championship standings computations.
- Deployed containerized API achieving 50ms response times and production reliability by implementing FastAPI with Docker, MLflow versioning, and automated health monitoring endpoints.

. **Credit Card Fraud Detection with Imbalanced Learning** (GitHub) — *Python, XGBoost, Scikit-learn, SMOTE*
- Developed XGBoost classifier on imbalanced data (0.17% fraud rate), applying SMOTE and stratified CV to achieve 0.94 recall, 0.88 precision.
- Engineered velocity features, z-scores, and entropy metrics, raising model AUC from 0.91 to 0.96 via domain-informed design.
- Achieved sub-200ms inference latency for practical real-time deployment in transaction authorization systems.