Thathsarani Sandarekha

Machine Learning Engineer

Email : thathsaranisandarekha@gmail.com

Portfolio , LinkedIn , GitHub, Medium

Professional Summary

Machine Learning Engineer with hands-on experience designing and deploying end-to-end ML systems for real-world applications. Strong background in applied machine learning, NLP, computer vision, and LLM-powered systems, with experience building production-ready APIs, data pipelines, and analytics workflows on cloud platforms. Proven ability to translate experimentation and research into scalable, maintainable solutions, supported by peer-reviewed publications in international conferences.

Technical Skills

Programming: Python, Java, SQL
Machine Learning: PyTorch, TensorFlow, Scikit-learn, Transformers, CNNs, LSTMs, Time Series Forecasting
NLP & LLMs: RAG, LangChain, BERT, LLaMA, SentenceTransformers
Computer Vision: YOLOv8, SAM, CLIP, OpenCV
Data & MLOps: Databricks, PySpark, Docker, MLflow, AWS, Azure
APIs & Apps: FastAPI, Streamlit, React.js
Analytics: Power BI (DAX, Power Query), Pandas, NumPy, Excel

Professional Experience

Machine Learning Engineer (Intern)
OCTAVE – John Keells Group *Jul 2023 – Jul 2024*

- Designed and deployed scalable data pipelines on Azure Databricks using PySpark to ingest and preprocess large marketing and operational datasets, supporting near real-time analytics in Power BI.
- Developed an LLM-powered retrieval-augmented generation (RAG) assistant enabling non-technical stakeholders to query sales and finance data using natural language, returning contextual insights and visual summaries.
- Collaborated with data science, engineering, and finance teams within the Cinnamon Hotels group to deliver ML-driven solutions for budgeting optimization, revenue forecasting, and operational efficiency.
- Applied MLOps best practices including model versioning, cloud-based training, and monitoring workflows to ensure reliability and reproducibility of deployed ML systems.
- Performed data validation, reconciliation, and KPI verification using Excel and SQL prior to dashboard refreshes, improving data accuracy and stakeholder trust.

Projects

Air Quality Data Assistant – End-to-End ML Agent System

- Built an AI-driven analytics agent to query and analyze multi-room air quality sensor data directly from raw NDJSON files without a database.
- Implemented schema normalization, dynamic Python code generation, execution, and visualization pipelines with FastAPI and React.
- Enabled natural language queries with automatic chart and table generation, simulating real-world IoT analytics workflows.

Tech Stack: Python, FastAPI, Matplotlib, Pandas, SentenceTransformers, Groq LLaMA3, React.js **Demo Video**

MetaFaces – Consistent Multi-Character Image Generation (Research)

- Designed a memory-guided diffusion pipeline based on SDXL to improve character consistency across multi-image visual storytelling tasks.
- Integrated attention injection, anchor-based feature reuse, and spatial segmentation using Grounding DINO + SAM.
- Published at FTC 2025 (Springer LNNS) with benchmark improvements over state-of-the-art methods (CLIP Score: 0.3619, DreamSim: 0.1753).

Tech Stack: Python, Transformers, Diffusion Models, OpenCV, Grounding DINO, SAM **Demo Video**

Dual-Stage Brain Tumor Segmentation & Classification

- Implemented a two-stage deep learning pipeline combining U-Net-based segmentation with CNN-based classification for MRI scans.
- Achieved Dice score of 0.89 for segmentation and 94.5% classification accuracy.
- Designed the pipeline to reflect clinical-style workflows including preprocessing, evaluation, and result interpretation.

Tech Stack: TensorFlow, Keras, OpenCV

Additional Projects

- ChatGroq: Multi-file conversational RAG chatbot
- Park Direct - Edge AI Parking System: Real-time license plate recognition using YOLOv8 and PaddleOCR, optimized for edge deployment. **Demo Video**
- ResuMatch AI: LLM-powered resume parsing and JD–Resume matching system. **Demo Video**
- Suspicious Activity Detection: Real-time surveillance system detecting violence and weapons.

Publications

- *MetaFaces: Dynamic Memory Networks for Consistent Character Representation in Multi-Image Visual Storytelling* — Accepted at the **FTC 2025**, *Munich, Germany* — Springer LNNS.
- *Systematic Review of Consistent Multi-Character Image Generation* — Accepted at the **ICMVA 2025**, *Melbourne, Australia* — SPIE Proceedings.
- *Comparative Analysis of Smoke and Fire Detection Systems* — Accepted at the **ICIVC 2024**, *Suzhou, China.*

Education

BSc (Hons) Artificial Intelligence and Data Science 2022 – 2025
Robert Gordon University *Scotland, UK*