Seyed Hasan Mir Hosseini

Data Scientist / Machine Learning Engineer

seyedhasan.mirhoseini1367@gmail.com +358 44 932 6220 Finland https://hasanportfolio.azurewebsites.net

Summary

Data Scientist with 2+ years of hands-on experience building and deploying predictive and deep-learning models (Transformers, CNN, LSTM) in Python. Experienced with large time-series datasets, feature engineering, and production deployment using Docker, CI/CD, and Azure ML across the full ML lifecycle.

Technical Skills

Languages

Persian Native English Professional Working Proficiency Finnish Beginner

Work Experience

Data Scientist (Intern)

University of Eastern Finland • Internship

Aug 2025 - Nov 2025

- Built a Transformer-based foundation model on large time-series data, pretrained with self-supervised learning (VICReg) and fine-tuned for classification.
- Developed the full PyTorch pipeline: preprocessing, augmentation, pretraining, fine-tuning, and evaluation.

Master's Thesis Researcher — Seizure Detection

Tampere University • Full-time

Sep 2024 - Jun 2025

- Designed and compared predictive deep-learning models (LSTM–Transformer, CNN–Transformer, and multimodal fusion) for time-series classification.
- Implemented end-to-end pipeline in PyTorch: preprocessing, spectrogram & time-series feature extraction, training, and model evaluation.

Data Analyst

Gamification Group, Tampere University • Part-time

Jan 2024 - Jan 2025

- Processed and analysed large experimental datasets: cleaning, preprocessing, statistical analysis, and visualisation.

- Co-authored peer-reviewed publications based on the research findings.

Education

Statistical Data Analytics

Tampere University

Sep 2023 - Jun 2025

Master's Degree

Business Administration (MBA)

University of Tehran

Sep 2013 - Jun 2016

Master's Degree

Mathematics

Damghan University

Sep 2008 - Jun 2013

Bachelor's Degree

Project Highlights

Kaggle — Applied Predictive Modelling

Competition-grade tabular and time-series prediction. Gradient-boosting ensembles (LightGBM, CatBoost, XGBoost), Random Forest, and Autoencoders. Feature engineering and cross-validated optimisation; evaluated with ROC-AUC, RMSE, and F1.

`Python` `LightGBM` `CatBoost` `XGBoost` `scikit-learn` `Pandas` `NumPy`

HealthCompass — Health Analytics Web App

Longitudinal biomarker trend analysis over time-series medical records. RAG assistant with hybrid BM25 + semantic retrieval and multi-LLM failover (Groq, Gemini).

`Django` `LangChain` `RAG` `ChromaDB` `Python` `PostgreSQL`

CasvuAI — Multi-Agent Content Pipeline

Multi-agent RAG pipeline orchestrating 6+ specialised agents for multilingual content generation, integrated with a Django web interface. Built during a hackathon.

`LangChain` `LangGraph` `RAG` `Django` `Python` `Groq`

HasanAI — Production ML Deployment Platform

Personal portfolio website with live ML model prediction demos, a RAG-powered AI assistant, and a technical blog. Deployed predictive ML models to production on Azure App Service with live prediction endpoints. Built a modular inference registry routing requests to Azure ML endpoints, with automatic local-model fallback.

`Django` `Azure ML` `Docker` `GitHub Actions` `PostgreSQL` `LangChain` `RAG` `Python`