Musawer Muradi

Junior Data Scientist | AI/ML Enthusiast

Email: musavir.muradi@gmail.com
Phone: +48 516 434 463
Location: Poznań, Poland (Poland Work Authorized)
GitHub: https://github.com/mswr1995

Professional Summary

Data Scientist and AI/ML Engineer with demonstrated experience in machine learning pipelines, data analysis, and cloud computing. Strong foundation in statistical analysis, predictive modelling, and data visualization with expanding expertise in Generative AI and Large Language Models. Proven ability to build end-to-end ML solutions and work with large datasets through academic projects and professional experience. Passionate about applying AI/ML to solve complex real-world problems across diverse domains including healthcare, cybersecurity, and business analytics. Authorized for full-time employment in Poland with no sponsorship requirements.

Technical Expertise

Machine Learning: scikit-learn, supervised/unsupervised learning, ensemble methods, model evaluation
Deep Learning: TensorFlow, PyTorch, Keras, CNN, RNN, LSTM networks, transfer learning
Statistical Analysis: Hypothesis testing, regression analysis, time series forecasting, A/B testing
Data Analysis & Visualization: pandas, NumPy, matplotlib, seaborn, ggplot2, exploratory data analysis
Feature Engineering: Data preprocessing, dimensionality reduction, feature selection techniques

Computer Vision: Image classification, object detection, medical image analysis, real-time video processing
Time Series Analysis: ARIMA, LSTM forecasting, seasonal decomposition, trend analysis, financial modeling
Generative AI & LLMs: Strong foundation in transformer architectures, attention mechanisms, Familiar with RAG systems, fine-tuning techniques, and text generation applications
Text Analytics: Sentiment analysis, document classification, information extraction

Languages: Python, R, SQL
Data Processing: Data pipeline development, ETL processes
Version Control: Git, collaborative development workflows
Documentation: Jupyter notebooks, technical reporting, reproducible research practices

AWS Services: Basic understanding of data lake architectures
Cloud Platforms: Google Cloud Platform, distributed computing concepts
Development & Deployment: FastAPI, REST APIs, Docker, containerization, virtual environments
Databases: MySQL, MongoDB

AI & Machine Learning Projects

DocAssist - AI Document Assistant (for WithSecure)

GitHub: https://github.com/mswr1995/doc-assist.git
Technologies: Python, FastAPI, ChromaDB (Vector DB), Ollama LLM, Docker, Pydantic, uvicorn, RAG Architecture, REST APIs

- Developed production-ready RAG system processing PDF/DOCX/TXT files with intelligent text chunking and vector storage
- Built FastAPI microservice with auto-generated OpenAPI docs, comprehensive error handling, and Docker containerization
- Implemented semantic search using ChromaDB vector database with local LLM inference for privacy-focused AI solutions

End-to-End Classification System (for WithSecure)

Technologies: Python, FastAPI, Docker, uvicorn, Makefile

- Built complete ML classification pipeline from data preprocessing to deployment
- Implemented RESTful API using FastAPI for model serving
- Containerized application with Docker for consistent deployment
- Automated build and deployment processes using Makefile

Medical Image Classification System

GitHub: https://github.com/mswr1995/Dental_Implant_Classification.git
Technologies: Python, TensorFlow, PyTorch, CNN, EfficientNet, DenseNet

- Developed deep learning models for clinical X-ray analysis (apical lesions, dental implant classification)
- Implemented transfer learning with EfficientNet and DenseNet architectures
- Achieved high accuracy through custom CNN design and hyperparameter optimization

Micro Gas Turbine Energy Analysis

GitHub: https://github.com/mswr1995/MGT-Energy-EDA.git
Technologies: R, Python, Linear Regression, RNN, LSTM, Keras, KerasTuner

- Predicted electrical power output using advanced regression models and RNN architectures
- Implemented LSTM networks with KerasTuner for automated hyperparameter optimization
- Achieved significant RMSE reduction through systematic model tuning and validation

Time Series Forecasting with Deep Learning

GitHub: https://github.com/mswr1995/stock-market-prediction-using-LinearRegression-and-LSTM.git
Technologies: Python, TensorFlow, LSTM, RNN, Keras

- Designed LSTM/RNN architectures for behavioral pattern analysis and financial forecasting
- Implemented GPU-optimized training workflows for sequential data processing

Computer Vision Behaviour Classification

GitHub: https://github.com/mswr1995/HumanBehaviourAnalysis.git
Technologies: Python, TensorFlow, VGG16, Transfer Learning

- Developed real-time video analysis system for human behaviour classification
- Implemented transfer learning on VGG16 with custom feature extraction layers
- Optimized for inference speed and deployed in edge computing environment

Global Land Temperature Analysis

GitHub: https://github.com/mswr1995/global-temperatures-analysis.git
Technologies: R, Python, Time Series Analysis, Statistical Modelling, EDA

- Built comprehensive data processing pipeline for global temperature analysis
- Applied statistical testing ($p < 2.2e-16$ significance) for climate trend validation
- Developed forecasting models for long-term climate prediction

Image Steganography in Spatial Domain: Comparative Analysis

GitHub: https://github.com/mswr1995/secure-image-steganography-using-LSB-and-double-XOR.git
Technologies: Python, MATLAB, OpenCV, Image Processing, Cryptography

- Implemented and compared LSB, PVD, and hybrid steganography methods for secure data embedding in digital images
- Evaluated performance metrics (PSNR, MSE, entropy) to analyse security vs. payload capacity trade-offs
- Applied Arnold scrambling and RLE compression to enhance security and embedding capacity

Professional Experience

Startup Technical Lead | egasi.uz (Classi Connect), Ankara

August 2022 – October 2024

- Led technical strategy for new e-commerce platform launch targeting automotive and real estate markets
- Conducted code audits and implemented development best practices for startup engineering team
- Built user behaviour analytics system to track engagement patterns and optimize user experience
- Implemented data-driven optimizations resulting in 30% user engagement improvement

Database Systems Developer | Kefayat Group Co., Afghanistan

November 2018 – September 2020

- Designed and implemented centralized database system for enterprise operations
- Built automated reporting workflows to support business intelligence and decision-making processes
- Optimized data architecture for efficient storage, retrieval, and analysis of operational data
- Collaborated with multiple departments to understand data requirements and system integration needs

Marketing Specialist (Part-time) | Monawar Hospital, Afghanistan

March 2017 - March 2018

- Managed digital marketing campaigns across social media platforms for healthcare services
- Developed content strategy to increase patient engagement and brand awareness
- Analysed campaign performance using digital analytics tools to optimize outreach effectiveness

---

Education

Master of Computer Engineering – Thesis Stage (Expected December 2025)

Kocaeli University, Turkey
Focus: AI and Machine Learning Applications, Data Science Natural Language Processing
Relevant Coursework: NLP, Deep Learning, Data Science, Machine Learning

Bachelor of Computer Science in Software Engineering

Kabul University, Afghanistan
Thesis: ISP Help Desk System – Design & Implementation

Erasmus Exchange Program

Adam Mickiewicz University, Poland
Specialized Coursework: Logic & Computation, Big Data & AI, Distributed Systems

---

Certifications

HarvardX Certifications

- **CS50 Data Science** - Full-stack data analysis, modelling, machine learning (October 2024)
- **CS50 AI with Python** - ML, NLP fundamentals, computer vision, transformers (April 2024)
- **CS50 Programming with Python** - Advanced Python, Flask development (December 2023)

Professional Certifications

- **TOEFL Certificate** - English Language Proficiency (ETS, August 2022)
- **Google Analytics** - Advanced analytics and data interpretation (Google, August 2022)
- **Plagiarism Certificate** - Academic integrity certification (Indiana University, 2020)

Additional Qualifications

Languages: English (Fluent), Turkish (Native), Persian (Native), Uzbek (Native), Pashto (Intermediate)

Volunteer Experience:
Erguvan Association for Integration of Immigrants, Turkey - Contributing to immigrant integration programs, demonstrating cross-cultural communication skills essential for international development teams.