Paris, Île-de-France, France
+33 7 53 40 20 05
meharoo261@gmail.com
LinkedIn

MAHR ALI ARSHAD

Artificial Intelligence Engineer

Professional Summary

Results-driven AI Engineer with 2 years of experience designing, developing, and deploying end-to-end AI/ML solutions in Generative AI, NLP, Big Data, and Computer Vision. Published work in Scientific Reports for advancing multilingual modeling in low resource languages.

Technical Skills

- **Programming Languages:** Python, C/C++, Java, SQL, JavaScript, HTML, CSS
- **Machine Learning / Deep learning / Image Processing:** Keras, TensorFlow, CUDA, PyTorch, SnnTorch, Torch, Scikit-learn, OpenCV, Pandas, Numpy
- **NLP / Generative AI:** spaCy, NLTK, LangChain, Transformers
- **Cloud / MLOps:** AWS, Docker, Airflow
- **Automation / Development Tools:** n8n, Git, Streamlit, Flask, FastAPI
- **Databases:** MySQL, PostgreSQL
- **Languages:** Urdu (Native), English (Fluent), French (A2 & Learning), Arabic (A2)

Experience

Oct 2023 – Jan 2025
Research Assistant – Pattern Recognition Lab @ PIEAS (Islamabad, Pakistan)
Supervisor: Dr. Ijaz Hussain

- **Hussain, I., Arshad, M.M.A., Cheema, A.N. *et al.*** Enhancing Urdu hate speech detection through differential transfer learning and adaptive loss functions. ***Sci Rep* 15, 37407 (2025)**
  Introduced NUHONS dataset (18k instances, manually annotated) and achieved SOTA results (91.49% F1-score)
- Developed a production-ready code generation chatbot using Q-LoRa fine-tuned **LLaMa2**, with full Flask backend and database integration.

July 2023 – Sep 2023
AI Intern – Systems Limited (Islamabad, Pakistan)
Advisor: Sana Salman

- Developed EDA web application using Flask enabling non-technical users to apply 5+ ML models (Decision tree, Random Forest, SVM, etc.) with automated visualization, reducing analysis time.

PROJECTS

Global Air Pollution Prediction: BitGrit Challenge
Technologies: Scikit-learn, LightGBM, KMeans

- Ranked 4th at Bitgrit International Competition in Public Leaderboard achieving 89.79% accuracy using LightGBM.
- Used K-Means for clustering geospatial regions and feature engineered location based (latitude and longitude) and temporal based features

Power Consumption Prediction: Daecon AI/Big Data Competition
Technologies: Scikit-learn, LightGBM, Random Forest, XGBoost, CatBoost

- Achieved 7.50% SMAPE for forecasting **hourly based power consumption** for **100 different buildings** from 25 Aug 2024 to 31 Aug 2024.

Multi-Class Thoracic Disease Classification
Technologies: Pytorch, Snntorch, scikit-learn,

- Developed a novel implementation of spiking neural network architecture achieving 67% faster training than Swin Transformer with comparable accuracy on NIH chest X-ray dataset (112,120 images, 14 diseases), outperforming CNN, ResNet, VGG16, and ViT baselines.

Social Media Post Automation
Technologies: n8n, OpenAI, Claude, NocoDB, Facebook/Instagram API handling
- Developed an automated workflow using n8n to automate Facebook/Instagram post generation for a travel agency
- Used Claude 3.5 and GPT 4o mini to generate ideas and auto publish, captions, post, and images after approval from user via Telegram bot

Multi-Class Tumor Classification
Technologies: Pytorch, Snntorch, Scikit-learn
- Achieved 91% Accuracy and 90% Precision, Recall, and F1-score using Spiking Neural Network, classifying three types of tumors to help provide patients with another opinion.

Medical Named Entity Recognition
Technologies: BERT, Tensorflow
- Fine-tuned BERT on the NCBI disease corpus and CDR dataset, to identify disease and chemical terms.

Education

- **EPITA**– MSc in Artificial Intelligence Systems **02/2025 – 02/2026**
- **Pakistan Institute of Engineering and Applied Sciences (PIEAS)** – BSc in Computer Science **09/2020 – 01/2025**