

# Alharth Alhaj Hussein

Informatics Engineer, Data Scientist, Computer Vision



## Profile

- Data Scientist with **10+ end to end Machine Learning projects** (TensorFlow, Scikit-Learn, OpenCV, Hugging Face Transformers, Pandas, NumPy, Matplotlib, Seaborn) and expertise in Python.
- Ranked in Kaggle competitions; solved 150+ programming challenges (Codeforces, UVA).

## Practical Experience

### Data Scientist, ML Engineer

[SHAI for AI](#), Jordan | March 2024 — August 2024

- **Toxic Comment Classification:** Built a multi-label toxic comment classifier using Bi-LSTM with Attention and pre-trained GloVe embeddings using TensorFlow, achieving a ROC AUC score of 0.977. Integrated a real-time GUI for interactive toxicity analysis across six categories.
- **Handwritten Digits Images Classification:** Trained a deep Convolutional Neural Network (CNN) by applying, data augmentation and Batch Normalization and other techniques, achieving 97% accuracy in classifying handwritten digit images.
- **Arabic Letters Audio Recognition (SVM, GUI):** Developed Support Vector Machine (SVM) using librosa and scikit-learn to classify Arabic letters from audio inputs with 63% accuracy on 1,400 samples. Implemented advanced audio feature extraction (MFCC, spectral, rhythm features) and integrated a real-time GUI for recording and analysis.
- **Diamond Price Prediction using Random Forest:** Engineered key features and fine-tuned a Random Forest model, reducing RMSE by 9.4% on 50K+ entries.
- **Music Genre Classification with Ensemble Stacking:** Improved F1-score by 13% using tuned ensemble (RF, SVC, KNN) using scikit-learn on 14K+ tracks.
- **Data Science Salaries EDA with Power BI:** Analyzed 5000+ data science salaries and Built 5+ interactive dashboards for.

## Education

### Bachelor's in Informatics Engineering

[Aleppo University](#), Aleppo | September 2020 — September 2025

- **GPA:** 83%
- **Relevant Coursework:** AI, Computer Vision, NLP, Deep Learning, Software Engineering.
- **Academic Project:**
  - **Coursera-Based Recommendation System** (Python, Django, SQL, HTML/CSS):
    - Trained a deep learning model (LSTM and Transformer architectures) to generate personalized suggestions using data we scraped from Coursera.
    - Developed full-stack web application with user-friendly interfaces for the recommendation system.
    - Turned the scraped data into a relational database and integrated the database (MySQL) with python backend for seamless interaction.

## Details:

📍 Aleppo, Syria  
☎ +963 991292874

✉ [alharth.alhaj.hussein@gmail.com](mailto:alharth.alhaj.hussein@gmail.com)

📅 **Date of Birth:** 5 Jun 2001

🇸🇾 **Nationality:** Syrian

## Links:

[My portfolio](#)

[LinkedIn](#)

[GitHub](#)

[Kaggle](#)

[Codeforces](#)

[Certificates](#)

## Languages:

- Arabic (Native)
- English (B2)

## Courses:

- **Machine Learning Specialization** (Coursera, Stanford/DeepLearning.AI)
- **Deep Learning Specialization** (Coursera, Stanford/DeepLearning.AI)
- **Generative AI** (Coursera, Stanford/DeepLearning.AI)

- **Human Motion Capture System** (Python, OpenCV, Mediapipe, YOLO, Depth anythingv2, PyQt6):
  - Developed a desktop application for real-time 2D and 3D human pose estimation from any camera or video source.
  - Integrated the "Depth Anything V2" model to accurately estimate 3D depth (Z-axis), enabling realistic motion capture for animation and game development.
  - Implemented functionality to stream motion data in real-time via WebSockets to other applications (e.g., Unity) or export it to JSON files.

## Technical Skills

- **Languages:** Python, C++, Java, SQL, HTML/CSS
- **Tools:** Git, VS Code, Jupiter Notebook, Kaggle and Google Colab notebook, MySQL Workbench, Microsoft Office.
- **Frameworks:** Django, FastAPI, TensorFlow, Scikit-learn, PyQt6
- **Models:** YOLO, MediaPipe, OpenPose, Depth Anythingv2, BERT, T5, GPT-2
- **Domains:** Machine Learning, Computer Vision, NLP, Data Analysis.
- **Data Preparation:** Data cleaning, feature engineering, feature scaling and normalization, handling imbalanced data, and data augmentation for model training.
- **Data Collection & Reporting:** Data collection processes and tools (Google forms, Kobo Toolbox), Regular data analysis and reporting, Data visualization and dashboards using Power BI, python packages (Pandas, Matplotlib, Seaborn).
- **Information Management:** Archiving and documentation, Technical support in data management and analysis.

## Volunteer Work:

More than one year of volunteering in [Blue Bits](#) Team, authored 30+ summarized lecture notes for 100+ collage students, improving study efficiency for exams.