

Riad Mohamad Kassem

riad.kassem17@gmail.com | +961 76 066 297 | LinkedIn
Beirut, Lebanon

TECHNICAL PROFILE

Senior Computer Science student with a foundation in Mechanical Engineering, specializing in Artificial Intelligence and Machine Learning systems. Combines computational modeling expertise with modern data science techniques to develop predictive and scalable solutions. Experienced in designing and deploying production-grade machine learning pipelines.

EDUCATION

AI Career Accelerator – Advanced AI Program 2025–Present
Advanced MLOps, Diffusion Models, LLM Frameworks, AI Automation, End-to-End Open-Source Model Development

TheAIEngineers – AI Program 2025–Present
Machine Learning, Deep Learning, Computer Vision, Large Language Models (LLMs)

Lebanese University (LU) 2023–Present
Bachelor of Science in Computer Science

Lebanese American University (LAU) 2021–2023
Completed two years of coursework toward a B.E. in Mechanical Engineering

SKILLS SUMMARY

Programming Languages: Python, Java, C, SQL
Libraries & Tools: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, OpenCV, Keras, Ultralytics (YOLO), HuggingFace Transformers, LangChain
Frameworks & Platforms: TensorFlow, PyTorch, Streamlit, Gradio, Jupyter Notebook, PyCharm, Visual Studio Code, IntelliJ IDEA, Git
Specialized Skills: NLP & Retrieval-Augmented Generation (RAG), Speech-to-Text (Whisper), Vector Databases (FAISS/ChromaDB), AI/ML Model Deployment
Soft Skills: Excellent Communication & Teamwork, Problem-solving & Adaptability, Time Management

LANGUAGES

Arabic: Native
English: Fluent

PROJECTS

RAG-Learn: Retrieval-Augmented Chatbot for Lecture Videos | GitHub | Live Demo

- Developed a Retrieval-Augmented Generation (RAG) assistant for querying lecture videos using LangChain, HuggingFace Transformers, and FAISS.
- Extracted transcripts from videos (English and Lebanese Arabic dialect) and built a vector database for efficient content retrieval.
- Integrated speech-to-text using Whisper for automated audio transcription, enabling multimodal input.
- Implemented a query interface allowing users to ask questions in English, with context-aware answers generated from video content.
- Built an end-to-end pipeline from data ingestion and embedding generation to retrieval and answer generation, showcasing advanced NLP and RAG deployment skills.

Custom Object Detection | GitHub

- Developed and deployed a YOLO-based object detection model on a **custom dataset annotated using Label Studio** (162 images).
- Structured and split the dataset into training and validation sets to ensure robust model evaluation.

- Trained the YOLO11s model using Ultralytics, achieving accurate object detection on validation images.
- Implemented real-time inference on images and live video streams, demonstrating end-to-end AI/ML pipeline expertise.

Safety Helmet Detection | [GitHub](#) | [Live Demo](#)

- Engineered a YOLO-based object detection pipeline for automatic safety helmet detection in images and live video streams.
- Curated and annotated a custom dataset, trained YOLO11s with Ultralytics, and achieved high detection accuracy under varied conditions.
- Delivered real-time inference with a Streamlit web interface, demonstrating full end-to-end AI/ML deployment.

Chest X-Ray Pneumonia Detection | [GitHub](#) | [Live Demo](#)

- Built and trained a Convolutional Neural Network (TensorFlow/Keras) to classify chest X-rays as Normal or Pneumonia.
- Applied preprocessing, normalization, and augmentation to improve model generalization on 5,863 images.
- Achieved 81% recall for pneumonia detection, improving sensitivity in medical image classification.
- Evaluated performance with accuracy, precision, recall, and confusion matrices, with visualizations for insights.

Credit Card Fraud Detection | [GitHub](#)

- Implemented an Isolation Forest model with PCA to detect anomalies in 284,807 credit card transactions.
- Addressed extreme class imbalance (0.17% fraud cases) through preprocessing and scaling.
- Achieved 25% recall on fraud cases, detecting unusual spending patterns in highly imbalanced data.
- Visualized results with PCA-transformed feature space and evaluated using classification reports and confusion matrices.

Heart Disease Prediction | [GitHub](#)

- Developed and compared Logistic Regression, Random Forest, and XGBoost for cardiovascular risk prediction.
- Achieved 98.83% precision with XGBoost, reducing false positives compared to traditional models.
- Built an end-to-end pipeline including feature engineering, missing value imputation, encoding, and hyperparameter tuning.
- Validated results with cross-validation and visualization to ensure robustness and generalization.

Car Price Prediction | [GitHub](#)

- Built a Ridge Regression model for car price prediction with preprocessing, feature engineering, and hyperparameter tuning.
- Achieved an R^2 score of 0.225, highlighting the limitations of linear models on complex, non-linear pricing data.
- Evaluated model performance and presented findings through visualizations.

CERTIFICATES

- | | |
|--------------------------------------------------------------------------------|-------------------------|
| • AI Engineer Explorer Certificate Course | School of AI, 2025 |
| • Large Language Models: Everything You Need to Know to Build Your First Agent | TheAIEngineers, 2025 |
| • Hands-On AI: Computer Vision Projects with Ultralytics and OpenCV | LinkedIn Learning, 2025 |
| • Deep Learning: Image Recognition | LinkedIn Learning, 2025 |
| • Computer Vision – From Image Basics to Real-World Applications | TheAIEngineers, 2025 |
| • Deep Learning: Everything you need to build your Neural Network! | TheAIEngineers, 2025 |
| • Machine Learning – Unsupervised Learning: Theory & Practice | TheAIEngineers, 2025 |
| • Machine Learning – Supervised Learning: Theory & Practice | TheAIEngineers, 2025 |
| • Python & Git: Programming Essentials | TheAIEngineers, 2025 |