Abdulrahman Ragab Mohamed

Artificial Intelligence Engineer

Location: Cairo, Egypt | phone: (+20) 1006210233 | E-mail: abdoragabsh10@gmail.com

LinkedIn: https://www.linkedin.com/in/abdulrahman-ragab-a35054248/ | GitHub | Youtube

Summary

AI Engineer with expertise in Python, ML, DL, NLP and LLMs, eager to build impactful AI technologies. Highly motivated to deliver data-driven solutions that contribute to business success.

Work Experience

AI Internship

YOUXEL | 14-July 2025 – 22-August 2025

- Developed complex, multi-step AI workflows in Flowise and LangChain to automate business processes, including task classification, and dynamic response generation using local open-source LLMs (e.g., Qwen, Phi).
- Implemented domain-specific AI assistants leveraging open-source language models like DeepSeek and Cohere for automated classification and analysis.
- Utilized natural language processing techniques with docling package in LangChain to enhance text parsing and classification capabilities. Designed modular AI workflows with structured JSON parsing and advanced prompt engineering techniques

Freelance Data Analyst

Self-employed | January 2023 – October 2023

- Ensured 100% data reliability by performing thorough data preprocessing, data cleaning and verification, guaranteeing accuracy, consistency, and completeness of all processed information.
- Enhanced client understanding by developing clear presentations and visual dashboards using Power Point and Power BI, effectively communicating key insights from complex datasets.

Education

Bachelor of Science in Computers and Artificial Intelligence | GPA: 3.46 (*Excellent)*

Major: Medical Informatics

Helwan University, Cairo, Egypt

August 2019 – June 2023

Projects

1- **Graduation Project: Fitness.ai (Grade: 96, A+)**

- Collected and preprocessed a custom dataset from various video sources, including collaborations with physiotherapists and I Developed an AI-driven application for real-time exercise classification and recognition using **MediaPipe** for human pose estimation and **Machine Learning models**, achieving **90.3%** accuracy
- Implemented a complete pipeline including video processing, landmark extraction (MediaPipe), manual labeling, and multinomial logistic regression for multi-class classification.
- Evaluated multiple models (ANN, Random Forest, Gradient Boosting) and finalized multinomial logistic regression for better generalization.

2- **Medical Insurance Complaints Review**

- Designed and implemented an AI-powered chatflow using Flowise and Cohere LLMs to automate medical insurance complaint triage, routing, and resolution for insured members.
- Developed a system leveraging prompt engineering and structured JSON parsing to classify complaints (basic/complex), delegate to appropriate medical levels, and manage coverage decisions (accept/reject), resulting in automated and dynamic user communication.

3- **Idea Project Analysis & Evaluation**

- Implemented a multi-stage workflow with JSON parsing and domain-specific analysis, creating a comprehensive evaluation framework that generates structured insights on idea feasibility, resources, approaches, and risks.
- Developed an AI-driven Idea Project Analysis & Evaluation system using Flowise and DeepSeek LLM, capable of automatically classifying and analyzing idea proposals across healthtech, quantum, and agentic domains.

4- **MLflow-Based MLOps Pipeline for Regression & Classification Tasks**

- Developed and tracked end-to-end ML workflows for both regression (**House Price Prediction**) and classification (**Iris Flower Classification**) tasks.
- Implemented hyperparameter tuning using GridSearchCV and tested multiple model configurations.
- Logged experiments with MLflow, including parameters, metrics, artifacts, and model versions and utilized the MLflow Model Registry for model management and lifecycle handling.

5- **RAG Document Q&A Assistant**

- Built RAG system using LangChain and OpenAI GPT-4o-mini to answer PDF queries with higher accuracy than base models. Engineered ChromaDB vector database with Hugging Face embeddings (all-MiniLM-L6-v2), reducing document search time. Developed Streamlit UI for source document inspection.

6- **Q&A Chatbot:** Developed a scalable, interactive chatbot application with Streamlit, enabling end users to ask questions and receive real-time answers powered by OpenAI GPT-4, GPT-4o-mini, and GPT-3.5-turbo models.

7- **IMDB Movie Review Sentiment Analysis**

- Processed and balanced a large dataset of labeled movie reviews with detailed NLP pipeline.
- Built and trained a **Bidirectional LSTM model** to classify IMDB movie reviews as positive or negative, achieving 85–90% test accuracy.
- Deployed a real-time sentiment analysis web app using Gradio, enabling user interaction and instant predictions.

8- **Email Message Classification**

- Built an NLP model to classify email messages (e.g., spam/ham) using TF-IDF, Bag-of-Words (BoW), and cleaned text using NLTK and Regex. Trained and evaluated Logistic Regression and Gradient Boosting models, achieving 92% accuracy (GB outperformed baseline).

9- **Medical Insurance Analysis & Prediction**

- Developed Gradient Boosting model predicting medical costs (RMSE: 6,788, R^2: 0.56) by analyzing demographics/health factors. Engineered data pipeline with feature standardization, one-hot encoding, and hyperparameter tuning via GridSearchCV. Evaluated 4+ algorithms (Random Forest, Polynomial Regression, etc.), achieving 10% lower error than baseline linear regression

10- **Intel Image Classification**

- Developed a deep learning model to classify images into 6 categories (buildings, forests, glaciers, mountains, sea, streets) using TensorFlow/Keras. Engineered a CNN architecture with multiple Conv2D, MaxPooling, Dropout, and Dense layers, achieving high accuracy on training data.

11- **Liver Patient Classification**

- Boosted Prediction Accuracy**:** Developed Gradient Boosting model achieving 79.5% accuracy in liver disease detection by analyzing clinical parameters, outperforming 7+ other ML algorithms. Engineered preprocessing workflow with missing value imputation, feature standardization, and PCA, optimizing model efficiency.

12- **Rock Density X-ray Classification**

- Achieved 0.09 RMSE (Gradient Boosting), beating 7 models (Random Forest: 0.13). Early stopping cut estimators by 68% (120→38) while keeping 95% accuracy. Automated evaluation with run_model() for instant RMSE+plots.

Skills

- **Programming Languages:** Python, Java, C.
- **Python Libraries & Frameworks:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, NLTK, SpaCy, MediaPipe, OpenCV, LangChain, Streamlit.
- **Tools & Platforms**: Version Control (Git), Jupyter Notebook, VS Code, Google Colab, Microsoft Excel.
- **Software Development:** Object-Oriented Programming (OOP), Data Structures & Algorithms
- **Software Engineering:** SDLC (Software Development Life Cycle).
- **Database Management:** Database Design, SQL Server, Query Optimization.
- **Data Analysis & Visualization:** Exploratory Data Analysis (EDA), Data Transformation, Data Wrangling, Data Cleaning, Preprocessing, Statistical Analysis, Feature Engineering, Power BI (Dashboards).
- **Mathematics for Data Science:** Calculus, Linear Algebra, Probability, Statistics.
- **Machine Learning:** Supervised Learning (Regression, Classification), Unsupervised Learning (Clustering, Dimensionality Reduction), Model Evaluation, A/B Testing, MLOps Concepts.
- **Deep Learning:** ANN, CNN, RNN, LSTM, GRU, Transformers.
- **Natural Language Processing (NLP):** Sentiment Analysis, Text Classification, Text Clustering.
- **Generative AI:** Large Language Models (LLMs) e.g. OpenAI, Groq, Hugging Face and Ollama open-source models, RAG (Retrieval-Augmented Generation), Chatbots, Prompt Engineering, Vector Database, AI Agents.

Certificates & Courses

- Machine Learning and Data Science Diploma
- Unsupervised Machine Learning Diploma
- Complete Generative AI Course (LangChain & Hugging Face)
- Complete AI Engineer Bootcamp
- Python for Data Science and Machine Learning Bootcamp
- TensorFlow for Deep Learning Bootcamp
- Statistics Fundamentals
- Become a Linear Algebra Master