Abdalla Abdalla

US Citizen +1 (612)-201-5656 abdalla7294@outlook.com Linkedin Github Website

Education

University of Cambridge

Bachelor's in Physics – Courses including: *Cambridge, United Kingdom*

- Quantum Mechanics, General Relativity, Machine Learning, Statistical Modeling, Probabilistic Graphical Models

Skills

Programming and Scripting Languages: Python, SQL, R, HTML, CSS, JavaScript, LaTeX, Markdown, MATLAB
Libraries and Frameworks: Django, TensorFlow, PyTorch, Scikit-learn, YOLO, XGBoost, OpenCV, Next.js, Matplotlib, PySpark
Data Engineering Tools: Apache Airflow, Apache Spark, Databricks, Azure Data Factory
Cloud Platforms: Google Cloud Platform (GCP), Microsoft Azure
DevOps and Automation: Docker, Kubernetes, Terraform, CI/CD, Git, Version Control, MLflow, Azure DevOps
Data Visualization and Business Intelligence Tools: Power BI, Looker, Tableau, Azure Synapse Studio

Experience

Data Analyst December 2024 – March 2025

Fairview Health Services *Minnesota, United States*

- Optimized workforce planning by 30% through UKG Pro Human Capital Management (HCM) suite, analyzing real-time employee scheduling data to ensure efficient resource allocation and compliance with hospital policies.
- Decreased last-minute absences by 36% by developing predictive staffing models using Python and SQL, allowing for proactive workforce forecasting and real-time shift adjustments based on historical trends.
- Reduced manual data entry errors and improved reporting efficiency by automating HR data pipelines, streamlining the generation and distribution of scheduling reports across multiple hospital units.

Software Engineering Fellowship July 2024 – September 2024

Headstarter *Remote, United States*

- Increased user engagement in training regimes by 25% through A/B testing, demonstrating the effectiveness of a new computer vision model, which provided precise feedback and actionable insights for technique improvement.
- Achieved 94% accuracy in tracking and evaluating fighters' movements by leading computer vision efforts with OpenCV, Roboflow, Mediapipe and YOLO, delivering precise visual feedback to enhance technique refinement.
- Reduced video processing time by 35% and improved model scalability by deploying machine learning models with Docker and Vertex AI, accelerating inference speeds and enabling real-time video analysis.

Data Scientist Intern July 2023 – September 2023

Thornton Tomasetti *Warrington, United Kingdom*

- Improved structural failure prediction accuracy by 22% with advanced machine learning models including Random Forests, Gradient Boosted Trees, and XGBoost, identifying critical factors such as material composition and load patterns.
- Enhanced catastrophic event risk assessments by 31% using Monte Carlo simulations and Latin Hypercube Sampling, which reinforced safety protocols and mitigated potential hazards to protect both personnel and infrastructure.
- Designed a cost-optimization framework, saving a client £84K, streamlining uncertainty quantification for data-driven decision-making.

Data Analyst Intern June 2019 – November 2019

Nuffield Foundation *Manchester, United Kingdom*

- Improved image resolution by 15% by designing a specialized satellite receiver with a "V" dipole antenna, which optimized the capture of NOAA satellite signals.
- Enhanced the signal-to-noise ratio of NOAA satellite imagery by 27% through refined frequency tuning and signal processing techniques, enabling more precise detection of environmental phenomena.
- Developed advanced GIS data workflows that integrated high-resolution satellite imagery with spatial datasets, streamlining emergency response planning and reducing incident response times.

Projects

PDF RAG App | *Python, Streamlit, PyPDF2, LangChain, Pinecone, Groq, SentenceTransformer* |

- Designed an AI-powered document retrieval system for enterprise knowledge management, improving search efficiency by 60% over traditional keyword-based methods.
- Integrated SentenceTransformer embeddings with Pinecone and Groq's LLM API for efficient Retrieval-Augmented Generation.

Tumor Track | *Python, NumPy, Matplotlib, pandas, seaborn, TensorFlow, OpenCV, scikit-learn* |

- Achieved 88% F1-score on MRI-based tumor classification using a Convolutional Neural Network (CNN), supporting early detection strategies in healthcare AI applications.