

Muhammad Abdulsalam

Data Scientist | Machine Learning Practitioner | AI Engineer

Egypt | mohameedmtmn@gmail.com | mohamed1249.github.io | github.com/mohamed1249 | linkedin.com/in/muhammad-abdulsalam-b127641b9 | kaggle.com/ma12492002/code

PROFILE

Data Scientist and AI Engineer with 3+ years of hands-on freelance and part-time experience delivering machine-learning, analytics, automation, and AI workflow solutions. Built AI systems at Say Hai, delivered Upwork client projects across multiple industries, and taught AI/Data Structures as a Teaching Assistant. Portfolio includes 60+ documented projects across churn prediction, forecasting, A/B testing, computer vision, NLP, RAG, LLM agents, APIs, BI dashboards, and reusable Python tooling.

TECHNICAL SKILLS

Data Science: EDA, data cleaning, feature engineering, statistical analysis, A/B testing, Bayesian testing, metric interpretation, reporting

Machine Learning: classification, regression, clustering, forecasting, XGBoost, Random Forest, SVC, Logistic Regression, LSTM, GRU, CNN, RNN

AI / NLP: RAG, LLM agents, document QA, embeddings, vector search, LangChain, LangFlow, Ollama, OpenAI API, Claude API, GPT-4, Llama

Visualization: Tableau, Power BI, Plotly, Matplotlib, Seaborn, dashboard design, stakeholder-readable reporting

Engineering: Python, SQL, Pandas, NumPy, Scikit-learn, TensorFlow/Keras, PyTorch, Flask, Streamlit, Gradio, Docker, GCP, Azure, MongoDB, Git

EXPERIENCE

Freelance Data Scientist & AI Engineer | Upwork / Independent Clients | 2022 - Present | Remote

- Delivered data science, machine learning, AI, and technical-writing projects for clients across the United States, Australia, Malaysia, UAE, Netherlands, China, Canada, and Europe.
- Built models for churn prediction, donor prediction, incident forecasting, property transaction analysis, clustering, classification, forecasting, and app-review scraping.
- Produced client-facing reports, executive summaries, code deliverables, dashboards, and reusable Python workflows.

AI Engineer (Part-Time) | Say Hai | Apr 2023 - Apr 2025 | Remote

- Developed custom AI workflows for customer organizations using OpenAI API, Claude API, Python, GCP, Flask, Gradio, Streamlit, LangFlow, LangChain, n8n, and Google APIs.
- Designed Flask API backends, deployed workflows on GCP, and connected LLM tools to client-facing web experiences.
- Processed and analyzed customer data to turn business needs into AI-powered solutions and workflow automation.

AI Teaching Assistant | Misr Higher Institute For Commerce & Computers | Sep 2024 - Feb 2025 | Egypt

- Taught Artificial Intelligence to senior Computer Science students and Data Structures to second-year students.
- Delivered lectures, explained complex concepts, supported hands-on coding exercises, and assisted with grading and mentoring.

SELECTED PROJECTS

Graduation Project: AI-Powered Mental Health Assistant | Python, OpenAI API, LangChain, Flask, Gradio, MySQL

- Built an LLM-powered assistant using DSM-5 and research-paper content to provide mental-health information and guidance around Autism and ADHD.
- Integrated Python, OpenAI API, Gradio, Flask, PHP, JavaScript, HTML, MySQL, and LangChain into an end-to-end application.

M-Ana: Python Package for ML and Data Science | Python, Packaging, Data Science Utilities, Documentation

- Designed a Python package with reusable functions, classes, and modules for machine-learning, data-science, and data-analysis workflows.
- Prepared GitHub documentation, reports, and publication support for developer-facing use.

Predictive Analytics for Work Orders | Python, Pandas, XGBoost, Random Forest, LSTM, KNN, ARIMA

- Built a time-series machine-learning solution to predict work-order loads and support forward-looking operational planning.
- Compared multiple models, tuned features and hyperparameters, and delivered dashboard-ready predictions with reported accuracy up to 96%.

Customer Event Prediction for Subscription Retention | Python, TensorFlow, LSTM, Scikit-learn

- Transformed about 119k raw subscription-event rows into a 52k-row modeling table with event, subscription, box, and timing features.
- Predicted each customer's next event and expected timing, reaching 85.9% event accuracy and 60.7% exact event-and-date match.

Telecom Customer Churn Prediction | Python, XGBoost, Scikit-learn, Power BI

- Built a churn pipeline from weekly telecom records, usage/revenue data, recharge behavior, demographics, and disconnected-number labels.
- Final XGBoost workflow reached 79.6% accuracy, 89.2% recall, 80.4% F1, and 80.2% ROC AUC.

Traditional vs. Bayesian A/B Testing | Python, SciPy, Bayesian Testing

- Compared retention outcomes for 90,189 Cookie Cats players using t-tests, chi-square tests, Beta posteriors, and Thompson Sampling.
- Recommended gate_30 based on stronger 7-day retention evidence and explained the difference between p-values and posterior probability.

Fully Open-Source RAG System | LangChain, Ollama, Llama 3, FAISS, Streamlit

- Built a local PDF question-answering app with document loading, chunking, embeddings, vector search, retrieval, and LLM answering.
- Demonstrated privacy-friendly RAG using open-source components rather than closed hosted model APIs.

EDUCATION

Master's Degree in Data Science, Cairo University - started Oct 2025 / in progress

Bachelor's Degree in Computer Science, Misr Higher Institute For Commerce & Computers - 2021 to 2024

Udacity Machine Learning Cross-Skilling - 2022 | Udacity Advanced Data Analysis - 2022

CERTIFICATIONS & LEARNING

Google Advanced Data Analytics; Kaggle courses; ITI Applied Deep Learning; Udacity Data Engineering Nanodegree; Udacity Advanced Data Analysis; Udacity Machine Learning Cross-Skilling.

LANGUAGES

Arabic: native | English: professional working proficiency