Omar Elkobrossy

859-420-0509 omarelkobrossy@yahoo.com linkedin.com/in/omar-elkobrossy github.com/omarelkobrossy

Education

Goldsmiths, University of London

B.S. Computer Science | Expected May 2028

- Relevant Coursework: Data Structures & Algorithms, Object-Oriented Programming, Software Design & Development, Discrete Mathematics, Linear Algebra, Probability & Statistics

Technical Skills

Languages: Python, SQL, C++, JavaScript, HTML/CSS
Machine Learning & Data: XGBoost, Scikit-learn, TensorFlow, Pandas, NumPy, Optuna, Time-Series Analysis, Feature Engineering, Walk-Forward Validation
Tools & Technologies: Git, Docker, Linux, Bash, AWS, REST APIs, SQL Server, Flask, FastAPI
Concepts: Machine Learning Pipelines, Real-Time Data Processing, Drift Detection, System Design, Data Structures & Algorithms

Experience

Software Engineer — Dec 2023 – Present

Bluegrass Integrated Communications — *Lexington, KY*

- Reduced customer and production data processing time by **46%** by building an event-driven pipeline that ingests and consolidates data from multiple internal sources.
- Automated document and layout generation by developing a placeholder-based templating system used across internal workflows.
- Engineered middleware to validate, enrich, and cross-reference customer submissions with internal production databases.
- Maintained operational continuity by building a GUI tool with manual override capabilities and error handling for failure scenarios.

Software Developer — Oct 2021 – Mar 2022

Vortex Academy — *Alexandria, EG*

- Built and programmed ROV mobility systems using C++ and Arduino, contributing to a first-place competition finish.
- Developed Python-based algorithms for technical and competitive programming challenges.
- Designed, tested, and iterated on ROV subsystems to improve reliability during competition scenarios.
- Implemented computer vision functionality to support object detection and navigation tasks.

Projects

End-to-End Adaptive Regime-Aware ML Trading System | *Python, XGBoost, AWS, Docker* | ***GitHub Repo***

- Built an end-to-end machine learning system for non-stationary time series by integrating **real-time inference**, periodic retraining, and live execution workflows.
- Improved training and inference consistency by designing a leakage-safe feature pipeline with **200+ features**, past-only transformations, and online statistical scaling.
- Adapted model behavior across changing market conditions by developing a **regime-aware** parameter tuning system (FAPT) using **Wasserstein-based similarity** between historical states.
- Implemented event-driven retraining via statistical **drift detection (PSI)** and **walk-forward validation** to maintain performance under shifting data distributions
- Engineered a scalable optimization and execution framework, combining joint model + strategy tuning (Optuna) with **distributed AWS Batch** processing

CheXpert Multi-Label Chest X-Ray Classifier | *Python, PyTorch, ConvNeXt*

- Built an end-to-end transfer learning pipeline on **223K+ chest X-ray images (440GB)** using **ConvNeXt** and **ImageNet-based** feature initialization.
- Reduced data leakage risk by implementing **patient-level stratified splitting** that preserves rare pathology distributions across train and validation sets.
- Improved robustness to class imbalance and uncertain medical labels by developing weighted loss strategies and soft-target label handling.
- Increased training visibility and debugging capability by implementing diagnostics for gradient health, calibration behavior, prediction collapse, and GPU utilization.
- Improved experiment continuity by designing resumable training workflows with separate model checkpointing and training-state recovery.