Mohamed Awadalla

Brooklyn, NY | 917-436-9873 | Mohamedawadalla75@gmail.com | github.com/Mo-Awadalla | moawadalla.com

Education

Long Island University-Honors College — Brooklyn, NY
Bachelor of Science in Computer Science — Expected May 2026
Honors: Dean's List, Dean Scholar

Skills

- **Languages & Frameworks:** Python, C++, JavaScript, TypeScript, HTML/CSS, SQL (PostgreSQL), PyTorch, Scikit-learn, NumPy, Pandas
- **Machine Learning & AI**: LLM Fine-Tuning (LoRA), Neural Networks, Transformers, NLP, Time-Series Modeling (XGBoost, Gradient Boosting), Backpropagation, Custom Autograd
- **Data Engineering:** Data Pipelines, Feature Engineering, Text Processing, Data Cleaning, OCR (Tesseract, AWS Textract), REST API Integration, ETL
- **Tools & Infrastructure**: Git, Linux, Docker, AWS, Cloudflare (Pages, Workers), Jupyter, REST APIs, CI/CD
- **Foundations**: Data Structures & Algorithms, Object-Oriented Design, Automatic Differentiation, Systems Design, Empirical Experimentation

Relevant Experience

New York City Emergency Management – Legal Tech & Automation Intern — June 2025 - Present | Brooklyn, NY

- Designed and built a custom event registration and tracking website (HTML, CSS, JavaScript) integrated with the Zoom Events REST API, handling sign-up and attendance tracking for 1,000+ registrants across hybrid events and replacing a manual coordination process with a reusable template adopted by the department.
- Authored a vendor evaluation framework benchmarking 8 competing legal matter management platforms (Thomson Reuters-class systems) on technical capabilities, cost, and integration fit; framework adopted by Chief Counsel and drove the department's final vendor selection.
- Designed and deployed an end-to-end OKR tracking and reporting system for the Office of the Chief Counsel: built a dynamic Power Apps form with cascading dropdowns backed by a SharePoint list of 63 pre-populated records across 4 attorneys, and orchestrated 3 Power Automate flows for quarterly reminders, 48-hour compliance follow-ups, and automated HTML report generation and email delivery.

Projects

Custom Autograd Engine & Character-Level Language Model | github.com/Mo-Awadalla/autograd, github.com/Mo-Awadalla/makemore

- Re-implemented a reverse-mode automatic differentiation engine in Python (inspired by Karpathy's micrograd), then extended it beyond the tutorial with tensor operations, a numerical gradient checker using finite differences, and a PyTorch benchmark suite validating gradient correctness against torch.autograd on identical inputs.
- Built a character-level language model on top of the custom engine with tokenization, embedding layers, and an MLP architecture following Bengio et al. (2003), trained end-to-end via backpropagation on a 32,000-name dataset.

LegalDocuMan — Document Processing & Classification Suite | github.com/Mo-Awadalla/LegalDocuMan

- Built a modular Python application for automated legal document classification, signature detection, and vendor-based file organization — deployed at NYC Emergency Management's Office of the Chief Counsel to process a legal drive of 16,000+ contracts ahead of migration to a new legal management system.
- Architected as a multi-component pipeline (processing engine, threaded GUI, CLI query tool, test suite) with OCR fallback via Tesseract, PDF text extraction via pdfplumber, and persistent metadata tracking for retention and destruction scheduling; released as MIT-licensed open-source with full installation and contribution documentation.

Stock Return Prediction (XGBoost & Gradient Boosting) | github.com/Mo-Awadalla/stock-return-prediction

- Built a Python pipeline (yfinance, XGBoost, Gradient Boosting) for next-day price prediction on AAPL using 9 engineered technical indicators; predicted next-day returns to handle non-stationarity, then reconstructed price for evaluation against a naive same-as-yesterday baseline on a held-out test set (5-year window, 70/15/15 chronological split). XGBoost achieved 1.54% MAPE vs. 1.05% for the baseline — a known finance-ML result where technical indicators capture price shape but not the news-driven variance dominating daily moves.
- Engineered time-series features across multi-API sources, ran controlled experiments on model architecture and training procedure, and analyzed performance across configurations to improve generalization on out-of-sample data.