

Ali Abouelazm

713-834-4423 | aliazm419@gmail.com | linkedin.com/in/ali-abouelazm | aliabouelazm.com | github.com/AliAbouelazm

Education

Texas A&M University

Aug. 2023 – May 2028

B.S. Data Engineering | Minor in Mathematics

College Station, TX

- Relevant Coursework: Data Structures & Algorithms, Quant Models for Machine Learning, Statistical Learning & Decisions, Optimization of Analytics, Computational Data Science, Principles of Data Science, Statistics

Experience

Tesla

Aug. 2026 – Present

Data Engineer Intern, Fleet Analytics

Palo Alto, CA

- Engineered and optimized **Airflow** DAGs for production data pipelines processing petabyte-scale vehicle telemetry using **Trino**, **Spark**, and **S3**, ensuring reliability across Fleet Analytics workflows
- Designed and deployed a live projection tool and executive dashboard tracking data storage migration from **AWS** to on-premise **VAST**, enabling real-time capacity planning across the engineering organization
- Built an **ML** model on raw **CAN bus** signals to automate vehicle fault diagnosis for technicians, implementing noise filtering pipelines to extract high-quality training signal from high-frequency sensor data

Cloudflare

May 2026 – Aug. 2026

Data Science Intern

Austin, TX

- Designed a **Narrative Mismatch Engine** in the **AEO** space, ingesting **3 years** of **SEMrush** data across **50K+** AI queries to detect and quantify gaps between AI-perceived and intended product positioning
- Applied **NLP** and **embedding-based** semantic analysis to surface narrative drift across **CDN**, **Zero Trust**, and **SASE** product lines, enabling content teams to close AI perception gaps at scale
- Automated query-level mismatch signal extraction from large-scale web and AI traffic data, increasing **AI content visibility** and driving measurable narrative alignment across product marketing surfaces

Texas A&M AgriLife

Mar. 2026 – Sept. 2026

Machine Learning Researcher

College Station, TX

- Trained **PyTorch**-based predictive models on **50GB+** of livestock biosensor data to forecast animal health indicators across **3+** active research deployments, iterating on architecture and feature engineering
- Architected scalable **ML** ingestion pipelines on **AWS** (S3, Athena) processing raw biological data streams, reducing preprocessing time by **35%** and accelerating iteration cycles for field researchers
- Automated feature engineering to extract health signals from noisy multi-sensor inputs, improving training throughput and prediction reliability across all production deployments

Projects

ExG Labs | Co-Founder | exglabs.com

Feb. 2026 – Present

- Co-founding a neural interface startup engineering a wrist-worn **sEMG** wearable that translates biosignals into device commands with near-zero false positives for motor-impaired users (ALS, stroke, spinal cord injury)
- Designed and implemented a **Probabilistic Fusion Layer** with confidence-gated intent resolution trained on **Ninapro DB5** sEMG data, incorporating cross-session **LDA** benchmarks and automated drift detection

Interview Mirror | PyTorch, FastAPI, React, Whisper, MediaPipe, WebSockets | mirror.aliabouelazm.com

May 2026

- Trained two emotion **CNNs** from scratch: a face model on **35K** FER+ images (**0.78** accuracy, **0.67** F1) with MixUp, CutMix, and EMA; a voice model on **9.4K** audio clips (**0.67** F1) with SpecAugment
- Engineered a **21-feature** fusion **MLP** combining face softmax, voice softmax, acoustic descriptors, and language signals into continuous confidence scores, deployed end-to-end at **500ms** via **WebSocket**

Drift | Python, FastAPI, React, DistilBERT, HuggingFace, SHAP | drift.aliabouelazm.com

Apr. 2026

- Fine-tuned **DistilBERT** on **40K+** labeled posts for 3-class sentiment classification, achieving **73.1%** F1 macro on a **10,360**-sample held-out test set across positive, negative, and neutral classes
- Designed an analytics engine computing rolling sentiment averages, anomaly detection with severity scoring, **TF-IDF** keyword extraction per class, and token-level **SHAP** explanations for model interpretability

Leadership

Lovable.dev

Jan. 2026 – Jul. 2026

Campus Ambassador

Texas A&M University

- Hosted AI development workshops for **200+** TAMU engineering students, driving campus-wide adoption of agentic AI tools

TAMU Datathon

Jan. 2026 – May 2026

Challenges Organizer

Texas A&M University

- Designed **ML** challenge problem sets for **500+** students spanning classification, NLP, and time-series tasks in Python and SQL

Technical Skills

ML & AI: PyTorch, scikit-learn, XGBoost, HuggingFace, LLM Tool Calling (OpenAI/Gemini APIs), MLflow, OpenCV, SHAP

Languages: Python, SQL, C/C++, TypeScript, JavaScript, R

Data & Cloud: AWS (S3, Athena), PySpark, Trino, Airflow, Databricks, Delta Lake, PostgreSQL, Docker, Git, FastAPI

Libraries: pandas, NumPy, HuggingFace Transformers, Matplotlib, Seaborn, SciPy