Ali Abouelazm

713-834-4423 | aliazm419@gmail.com | linkedin | website

Education

Texas A&M University **Aug. 2023 – May 2027**

Bachelor of Science in Data Engineering | Minor in Mathematics and Statistics *College Station, TX*

- Relevant Coursework: Linear Algebra, Database Systems, Data Structures Algorithms, Quant Models for Machine Learning, Optimization of Analytics, Computational Data Science, Differential Equations, Principles of Data Science.

Experience

TCG Digital Solutions **May 2025 – Aug. 2025**

Data Science Intern *Somerset, NJ*

- Architected an end-to-end **CV pipeline** using **Gemini Pro** and **OpenCV** to automate soccer highlight extraction; achieved **95%** precision across 100+ match hours and reduced manual processing time by **80%**.
- Developed and tuned gradient-boosting models (**XGBoost, CatBoost**) on large-scale tracking data to predict match outcomes; utilized **SHAP** values to provide tactical insights into player performance and xG.
- Implemented automated video alignment modules for kickoff and halftime detection, reducing data latency and improving highlight accuracy for downstream analysts.

Texas A&M University, Center for Teaching Excellence **Aug. 2025 – Dec. 2025**

Student Tech *College Station, TX*

- Engineered automated **ETL** pipelines using **AWS S3**, **Athena**, and **Python** to process institutional-scale datasets; reduced reporting latency for departmental leadership by **40%**.
- Designed production-grade BI dashboards in **QuickSight** to visualize faculty engagement and student outcomes, directly informing strategic resource allocation for **15+** departments.

Texas A&M University, ENGR 102 **Aug. 2025 – Dec. 2025**

Peer Teacher *College Station, TX*

- Mentored **100+** first-year engineering students in **Python**-based problem-solving, focusing on data structures, algorithmic logic, and collaborative debugging in weekly lab sessions.

Projects

Casual Marketing Impact | *Python, DoubleML, Streamlit, Docker, GitHub Actions* **Dec. 2025**

- Architected an end-to-end application utilizing **Double Machine Learning (DML)** to isolate the causal impact of marketing spend, controlling for **20+** confounding variables to eliminate selection bias.
- Engineered a production **MLOps** pipeline with **Docker** and **GitHub Actions** to deploy an interactive **Streamlit** dashboard, automating model validation and reducing manual reporting time by **90%**.

clinix.ai | *Python, SQL, scikit-learn, OpenAI API, Streamlit* **June 2025**

- Developed a **RAG-based** symptom extraction pipeline using the **OpenAI API** and **PostgreSQL**; achieved a **0.92 F1-score** on patient risk classification.
- Architected **SQL**-based clinical feature pipelines to process patient data, utilizing **Streamlit** to visualize real-time triage patterns for medical staff.

PL Predictor | *Python, pandas, NumPy, scikit-learn, BeautifulSoup* **June 2024**

- Engineered a predictive modeling framework using **XGBoost** to forecast league standings; implemented **30+** custom features including ELO-based strength metrics.
- Automated match data ingestion using **BeautifulSoup** and **Selenium**, ensuring **99%** data integrity through robust error-handling and validation scripts.

Leadership / Extracurricular

Tamu Datathon **March 2025 – Present**

Challenges Organizer *Texas A&M University*

- Directed the creation of Data Science challenges for **500+** students, utilizing **Python**, **SQL**, and **scikit-learn** to develop approachable yet rigorous ML problem sets.
- Provided real-time technical guidance and debugging support for participants during large-scale competitive events, ensuring high-quality submission outcomes.

Technical Skills

Languages: Python (Expert), SQL (Advanced), R, C/C++, JavaScript, TypeScript.
Machine Learning: PyTorch, TensorFlow, XGBoost, LLMs (OpenAI API), Causal Inference, A/B Testing, Scikit-learn.
Data & Cloud: AWS (S3, Athena, QuickSight), PostgreSQL, Snowflake, Docker, Git, Airflow.
Libraries: pandas, NumPy, Matplotlib, Seaborn, Plotly, Dask, SciPy, Hugging Face.