

# MAZEN EMAD ABDELHAIM

📍 Elgiza, Elgiza, Egypt ✉ Mazen7ektr@gmail.com 📞 +201011441742 🌐 in/mazen-emad14 🌐 mazen-emad-portfolio.netlify.app/

## SUMMARY

Final-year AI student at Benha University specializing in Data Science, engineering end-to-end Machine Learning, Deep Learning, and NLP pipelines, deploying scalable AI applications and dashboards utilizing Python, TensorFlow, and Streamlit.

## EXPERIENCE

### Data Scientist

X\_Clan

July 2025 - October 2025, Egypt-Mansoura

- **Processed** 500+ GB of raw data and engineered 40+ features, reducing data preprocessing time by 30%.
- **Deployed** machine learning models into production environments, successfully serving 10,000+ daily requests during real-world testing.
- Implemented predictive analytics pipelines utilizing Python and Apache Spark over 12-month project, forecasting customer behavior with 96% accuracy and improving acquisition conversion rates by 18% across 5 digital channels.

## PROJECT

### AI-Powered Meeting Transcriber & Summarizer

NTI • [github.com/MazenEmad14](https://github.com/MazenEmad14) • December 2025 - January 2026

- **Engineered** an automated audio processing pipeline utilizing Whisper ASR, transcribing **100+** hours of meeting recordings with **98%** accuracy.
- **Developed** autonomous CrewAI and Google Gemini agents to extract key action items, reducing manual documentation time by **85%**.
- **Integrated** automated summarization and follow-up distribution via Gmail SMTP, accelerating post-meeting communications for **50+** weekly stakeholders.

### AI-Powered Smart Attendance System

College • [github.com/MazenEmad14](https://github.com/MazenEmad14) • November 2025 - January 2026

- **Architected** a real-time facial recognition attendance system, processing live camera feeds at **30 FPS** with **99%** identification accuracy for **500+** personnel.
- **Engineered** secure backend pipelines to encrypt and manage **10,000+** daily attendance logs, eliminating manual data entry errors entirely.
- **Deployed** an administrative control dashboard, automating report generation and reducing attendance tracking overhead by **80%**.

### Credit Score Classification

College • [github.com/MazenEmad14/weatherAUS](https://github.com/MazenEmad14/weatherAUS) • December 2025 - December 2025

- **Engineered** a credit score classification model for 100,000+ customer records, achieving 95% predictive accuracy across three risk tiers.
- **Deployed** the final model via a Streamlit application, processing 5,000+ real-time predictions daily with sub-second latency.

## EDUCATION

### B.Sc. in Artificial Intelligence

BENHA UNIVERSITY - FACULTY OF COMPUTERS AND ARTIFICIAL INTELLIGENCE • Banha, Egypt • 2022-2026 • 3.5

## CERTIFICATIONS

### Natural Language Processing (NLP)

NTI • 2025

- **Implemented** comprehensive NLP pipelines utilizing text tokenization, TF-IDF vectorization, and dense word embeddings (e.g., Word2Vec/GloVe) to preprocess, clean, and classify **10,000+** text documents efficiently. **Engineered** text classification models using TensorFlow/Keras, benchmarking traditional machine learning algorithms against advanced deep learning architectures (RNN, GRU, LSTM), achieving an **88%+** validation accuracy and robust F1-scores. **Integrated** and fine-tuned pre-trained Transformer models (such as BERT) via the Hugging Face ecosystem to execute complex downstream tasks, significantly improving contextual language understanding.

### Machine Learning for Data Analysis

NTI • 2025

- **Executed** comprehensive Exploratory Data Analysis (EDA) on datasets containing **20,000+** records, utilizing Pandas, Matplotlib, and Seaborn to handle missing values, engineer predictive features, and identify core statistical correlations. **Trained** and optimized a diverse suite of supervised (e.g., Random Forest, SVM) and unsupervised (e.g., K-Means, PCA) machine learning models, alongside deep neural networks, achieving **90%+** test accuracy through rigorous hyperparameter tuning. **Developed** and deployed interactive web applications utilizing Streamlit, enabling users to input real-time data, execute dynamic model inference, and visualize complex dataset trends through intuitive dashboards.

### Artificial Intelligence Fundamentals

ITI • 2025

- **Applied** core mathematical principles—including linear algebra, probability, and advanced statistics—to formulate machine learning algorithms, utilizing numerical optimization techniques (e.g., Gradient Descent) to minimize loss functions effectively. **Executed** rigorous data preprocessing and Exploratory Data Analysis (EDA) pipelines on complex datasets of **10,000+** records, applying statistical methods to handle outliers, normalize distributions, and engineer robust features. **Architected** and trained Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) for complex pattern recognition tasks, optimizing model weights and achieving **85%+** validation accuracy on test datasets.

## SKILLS

Python, Pandas, Matplotlib, Seaborn, Scikit-Learn, XGBoost, Streamlit

TensorFlow, Keras

Natural Language Processing (NLP), NLTK, Text Processing, Sentiment Analysis, Language Modeling

Data Cleaning, Data Visualization, Feature Engineering, Model Development, Hyperparameter Tuning, Supervised Unsupervised Learning, Model Evaluation & Deep Learning  
DNN, CNN, Keras, TensorFlow

Data Structures, Algorithms, Problem Solving, Critical Thinking, Communication