

# Kareem Ahmed

[LinkedIn](#) | [Kaggle](#) | [Github](#) | [karimx791@gmail.com](mailto:karimx791@gmail.com) | +20 11 5347 2095

## EDUCATION

---

### Cairo University

OCT 2020 – JUL 2024

B.Sc. in Computer Science, Major in Artificial Intelligence, GPA: 3.03/4.0

## EXPERIENCE

---

### Data Science Vice Head at IEEE - Extracurricular Activities

SEP 2023 – DEC 2023

My responsibilities include educating members EDA, Data preparation, feature engineering, machine learning algorithms (I made an educational repository for ML algorithms and discussing pros and cons for each algorithm [\[Link\]](#)), and model deployment.

### Research Scientist & Engineer Intern at Owais Capital – Remote

OCT 2023 – NOV 2023

- Fine-tuning large language models (LLMs) to adapt them to Finance domain.
- Balancing between model performance and system constraints to deliver robust solutions.

### Data Science Diploma at Orange Digital Center – Cairo

SEP 2022 – NOV 2022

- I learnt scrapping, data preparation, machine learning algorithms, deep learning, NLP basics, model deployment.
- I applied every session task on the topics I mentioned before and made a final project to show skills that I learnt and deployed the model using flask, the project title is **flight delay prediction** [\[Link\]](#).

## PROJECTS

---

### Graduation Project

- Our graduation project involves developing a **live video translation system**, translating from English to Arabic and vice versa, with lip-syncing and maintaining the original speaker's voice.
- **Contribution:**
  - **Speech-to-Text Transcription:** Searched and used the best and fastest model for transcribing voice to text.
  - **Translation:** Identified and utilized the best and fastest model for translation from English to Arabic. As there was no existing dataset to test models, I used a dataset to evaluate model latency and performance. Fine-tuned the model to translate from the Egyptian dialect to English and deployed it on Hugging Face [\[Link\]](#).

### Image Multiclass Classification [\[Link\]](#)

- Used dataset on Kaggle with 10 classes each class describe the state of driver whether he is talking at phone, drinking, and other classes.
- Trained YOLO v8 on this dataset and I got 98% f1-score.

### Image Multiclass Classification [\[Link\]](#)

- Using CNNs only and with CNNs plus data augmentation to see its effect on accuracy.
- Using VGG16 only and with VGG16 plus fine-tuning to see its effect on accuracy.
- The best accuracy was achieved using VGG16 with fine-tuning, resulting in an F1 score of 91%.

### Health Insurance Cross Sell Prediction [\[Link\]](#)

- The main issue here is that the data is imbalanced, so I tried to handle this by oversampling and undersampling.
- The F1 score doesn't differ significantly even when using oversampling or undersampling, resulting in an F1 score of 81%.

## Competitions

---

### Market Type Classification Competition

Ranked 3rd (tied with 1st and 2nd) with a score of 98.9%, achieved solely through data analysis without machine learning models on a dataset with 3 classes.

### Vehicle Recognition Competition

Ranked 3rd with a score of 94.2%, using CNNs and pretrained models on a complex vehicle dataset with 10 classes.

## SKILLS

---

**Programming Languages:** Python, C++.

**Data Management:** SQL, MySQL, Data Cleaning, Data Preparation.

**ML & DL:** Machine Learning Algorithms, Deep Learning, NLP, CV, Keras, Scikit-learn, Pytorch.

**Version Control:** Git / Github