

# Ramzy Bakir

[ramzybakir.rsb@gmail.com](mailto:ramzybakir.rsb@gmail.com) | 01278673139 | Heliopolis, Cairo, Egypt  
[www.linkedin.com/in/ramzy-bakir](https://www.linkedin.com/in/ramzy-bakir) | <https://ramzybakir.github.io> | <https://github.com/RamzyBakir> | [www.kaggle.com/ramzybakir](https://www.kaggle.com/ramzybakir)

## WORK EXPERIENCE

- Converged Data Solution Architect Intern- Orange Egypt**; Smart Village, Egypt August 2024 - September 2024
- Collaborated with senior **solution architects**, understanding the structure of the **data centers** running **virtual machines**, maintaining, optimizing and deploying **docker containers** and **Kubernetes**.
  - Processed and wrangled **1 million records** representing **30k** users' internet usage sessions.
  - Developed and presented a dashboard featuring **24 KPIs** and **24 graphs**, providing actionable insights for bucket and offer recommendations.
- Data Analysis & Business Intelligence Intern - Centro**; Cairo, Egypt September 2023 - October 2023
- Gained practical expertise in leveraging **statistical analysis** and **data visualization** tools, including **Excel**, **PowerBI**, and **Incorta**, to drive data-driven insights and optimize decision-making processes.
  - Selected as one of the **top 3 interns** out of 12 for exceptional performance, particularly recognized for delivering the **best Power BI project**.
- RPA Business Analyst Intern - Advansys ESC**; Cairo, Egypt July 2023 - August 2023
- Developed expertise in **RPA business analysis**, with a focus on writing **PDDs**, creating **documentation**, and gathering client requirements.
  - Created Process Definition Documents (**PDDs**) for approved automation projects.
  - Collaborated closely with **clients** and the **development/solution** team to ensure balanced and successful project outcomes.
  - Initiated **two automation processes**, effectively managing **client meetings**, gathering requirements, conducting **business analysis**.

## EDUCATION

- EUI - Egypt University of Informatics** November 2022 - Present  
New Administrative Capital, Egypt
- B.S. Computing and Information Sciences**
- Relevant courses:** Algorithms Design and Analysis, Advanced Programming, Introduction to Artificial Intelligence, Introduction to Database Systems, Introduction to Software Engineering, Data Structures, Computer Architecture and Organization, Introduction to Probability and Statistics, Computational Linear Algebra, Programming Fundamentals, Introduction to Computing and Programming.

### Extracurricular Activities

- AI head - GDSC EUI** October 2024 – Present

## PROJECTS & PUBLICATIONS

- Ancient Egyptian NLP System** April 2025 - May 2025  
*Generative & Classificatory Modeling of Hieroglyphic Transcriptions* Python PyTorch, HuggingFace Transformers
- AegyptusTranslit1-gpt2-17M**
    - Designed and trained a **17M parameter GPT-2 architecture based model** from scratch for generating Ancient Egyptian transliterations.
    - Using a dataset of **~30,000** sentences of transliterated Ancient Egyptian; built a **BPE tokenizer** with a vocabulary size of **6,475 tokens**.
    - Implemented and optimized full **training pipeline** including data preparation, model training, evaluation, and text sampling.
    - View and use the model on **HuggingFace** via: <https://huggingface.co/RamzyBakir/AegyptusTranslit1-gpt2-17M>
  - roberta-aegyptustranslit-classifier**
    - Fine-tuned a RoBERTa-base** for **5-era** historical classification (Old Kingdom to Greco-Roman).
    - Optimized for limited data: the model is trained on a **10,000** balanced samples dataset, with AdamW (lr=2e-5, bs=64), for 20 epochs.
    - Model Achieved a **macro/weighted F1 score** of **~0.60** with a highest class F1 score of **~0.7**, lowest of **~0.4**.
    - View and use the model on **HuggingFace** via: <https://huggingface.co/RamzyBakir/roberta-aegyptustranslit-classifier>

### SYRA AGENT

- LLMs, GenAI* September 2024 - January 2025  
Python
- An **AI-driven resume screening pipeline** leveraging **Gemini 1.5's long-context processing** to handle **up to 2M tokens** per request.
  - Developed a **pipeline** to analyze and compare detailed resumes with job descriptions, providing **real-time** candidate fit assessments.

### Career Navigator Chatbot

- Functional programming, Machine Learning, Web scrapping, Web App* April 2024-May 2024  
Scala, Python
- Developed a **Scala-based backend** with a **Python frontend**, creating an **AI-driven career guidance platform** for job seekers.
  - Built an **NLP-powered chatbot** that **analyzes user queries** and provides **personalized job insights**, leveraging **keyword mapping and rule-based logic**.
  - Recognized by a **company department head**, leading to a successful **internship opportunity**.

### Egypt's Net Migration Patterns: A Time Series Analysis

- Study - Machine Learning Model* March 2024-May 2024  
Python
- This research aims to analyze how sociopolitical and economic factors have influenced Egypt's net migration over the past 50 years.
  - Was responsible for **data cleaning and transformation**, **exploratory data analysis (EDA)**, model selection, **model tuning**, and all **data visualization** to ensure robust and accurate **predictions**.
  - Developed and fine-tuned regression models, leveraging **Gradient Boosting**, **XGBoost** and conducted **feature engineering** to predict net migration.

## SKILLS AND CERTIFICATES

- Programming Languages:** Python, R, C, C++, SQL, Scala
- Data Visualization:** Matplotlib, Seaborn, Power BI, Tableau, Incorta, ggplot2, Tidyverse
- Data Cleaning & Transformation:** Pandas, NumPy, Beautiful Soup
- ML Algorithms:** Time Series Analysis, Regression Analysis, Gradient Boosting, KNN, Decision Trees, CNNs, Neural Network, Transformers
- ML Frameworks & Tools:** Scikit-learn, XGBoost, TensorFlow, PyTorch, LightGBM, CatBoost, Optuna, Hugging Face, LangChain, Streamlit
- LLMs & NLP:** Experience with developing, building and finetuning transformer architecture based model, LLM serving, agentic workflows and RAG.
- Statistical Analysis:** SciPy, A/B Testing, Hypothesis Testing, Bayesian Methods
- SWE & Automation:** UiPath, Microsoft Power Automate, Git, Jira, Docker
- Languages:** Arabic (Native), English & French (Fluent), German (Conversational)
- Certificates:** [Duke University Open Source LLMOps Solutions](#) | [UiPath Academy Automation Business Analyst](#) | [IBM Data Science Specialization](#) | [Data Scientist with Python - Datacamp](#) | [Machine Learning with Python - IBM](#) | [Data Analyst with Python - Datacamp](#) | [Data Analysis with R - IBM](#)
- ❖ You can access all the aforementioned project details, including source code and documentation, within my GitHub repositories via the following link:  
<https://github.com/RamzyBakir?tab=repositories>