Youssef Bastawisy youssefmahmod01@gmail.com

Junior Data Scientist www.linkedin.com/in/youssef-bastawisy

(+20)1207768339 / Egypt https://github.com/YoussefBastawisy

Professional experience

- **Applied AI Engineer** 08/2025 - Present
  Cycls
  Engineered sophisticated AI agents by leveraging Large Language Models (LLMs) with a focus on Retrieval-Augmented Generation (RAG) to ground outputs in factual data and reduce model hallucination. I extended agent functionalities by implementing function calling and tool-use protocols to enable real-time programmatic actions and interaction with external data sources. Additionally, I designed, built, and maintained the scalable and reliable infrastructure supporting these agents in a production environment.
- **Data Science Instructor** 08/2025 - Present
  Arabian Academy
  Empowered aspiring data scientists by delivering hands-on training in Python, data analysis, and machine learning. Designed real-world projects to build industry-ready skills and inspire data-driven problem-solving.
- **Machine Learning Intern** 07/2024 – 09/2024
  ZA Tech
  As a Machine Learning Intern at ZA Tech, I worked on designing and evaluating predictive models to support insurance and fintech solutions. My responsibilities included preprocessing real-world datasets, applying supervised and unsupervised learning techniques, and fine-tuning model performance using cross-validation and hyperparameter optimization. I also contributed to deploying models in a scalable environment using Python and cloud-based tools, while collaborating closely with data scientists and product teams to align technical work with business objectives.
- **Data Science Intern** 01/2023 - 2024
  WorldQuant University
  During my internship with WorldQuant University, I participated in hands-on projects that deepened my understanding of statistical modeling, data analysis, and machine learning. I worked on end-to-end pipelines involving data cleaning, feature engineering, model development, and evaluation using Python libraries such as Pandas, Scikit-learn, and TensorFlow. I also gained experience in working with real-world datasets across finance and health domains, and presented insights and results through interactive dashboards and written reports.
- **Data Science Instructor** 09/2022 – 03/2023
  PES - Programmers Elite School
  As a Data Science Instructor at PES, I was responsible for teaching foundational concepts in Python programming, data analysis, and machine learning to students and early-career professionals. I designed lesson plans, delivered live coding sessions, and guided students through practical projects involving real datasets. I also provided personalized feedback, helped troubleshoot code, and ensured students developed both technical and problem-solving skills aligned with industry standards.

Technical Skills

- **Programming Languages:** Python, R, SQL
- **Machine Learning & AI Tools:** Regression, Decision making, SVM, Logistics, Classification, K-Mean, KNN
- **Data Visualization Tools:** Matplotlib, Seaborn, Tableau, Power BI, Looker Studio, Excel
- **Data Analysis & Tools:** Pandas, NumPy, Excel, Power Query, SAS, Plotly

- **Deep Learning:** NLTK, SpaCy, CNN, LSTMs, RNN, TensorFlow, Keras
- **Database Management:** SQL Server, Oracle, MongoDB, Firebase, Dynamo DB
- **General:** Cloud, Data Modeling & Mining, Statistics, Probability Presentation, Communication

Projects Highlights

Graduation Project – Mind Care **09/2024 – 07/2024**
(NLP models– Firebase - Google Maps API, – VR development – Visualization - Data Processing & Analysis)

- Designed and developed a full-stack multi-platform system combining **AI**, **VR**, and **Mobile Health** to support over **100+ Alzheimer's patients and caregivers** during pilot trials.
- Created an **AI-powered Avatar** for cognitive assessment using **Natural Language Processing (NLP)** and **Speech Recognition**, achieving **94% accuracy** in interpreting MMSE responses across **500+ test cases**.
- Engineered a **VR-based cognitive training system** featuring adaptive difficulty and task personalization, resulting in a **35% boost in cognitive engagement scores** among users aged 60+.
- Developed a mobile application using **Flutter**, integrated with **Firebase**, **Google Maps API**, and real-time alerts to enhance caregiver-patient communication.
- Implemented **real-time location tracking**, which reduced simulated emergency response times by **up to 40%** across 10 test scenarios.
- Trained and fine-tuned **NLP models** on custom datasets of speech and behavioral patterns, reaching over **92% model accuracy** in early detection classification.
- **Awarded 1st Place** in the Faculty of Computer science & Artificial intelligence graduation project discussion

Sentiment Analysis on TechCrunch News Articles **10/2024 – 12/2024**
(BeautifulSoup – Feature Extraction – PCA for dimensionality reduction – clustering)

- Collected and processed **1,000+ articles** from *TechCrunch* using **BeautifulSoup** and **NewsAPI**, targeting technology-related news for sentiment evaluation.
- Cleaned and preprocessed text data using **NLTK** by applying tokenization, stop word removal, and lemmatization techniques.
- Extracted key terms using **TF-IDF**, highlighting top features and their relevance to sentiment classification.
- Visualized word importance using **WordClouds** and TF-IDF score distributions to gain insights into frequently discussed topics.
- Applied **KMeans clustering** with **PCA** dimensionality reduction during Exploratory Data Analysis (EDA) to detect thematic groupings among articles.
- Developed and trained a sentiment analysis model with a final **accuracy of 96%**, effectively identifying positive and negative news tones.

Certificates

- IBM Data Science Specialization
- Machine Learning Using SAS Viya 3.5
- Intermediate SQL Queries
- Data Science & Advanced Analytics Virtual Experience Program

Education

- Bachelor of Computer Science and Artificial Intelligence - Major(Data Science)
  Pharos University in Alexandria
  Grade: 3.8 GPA