

Mohamed Shehata

Data Scientist

✉ mohamed.elsayed.shehata@outlook.com ☎ +201143030697 📍 Alexandria, Egypt 🔗 LinkedIn 🐙 GitHub

EDUCATION

9-Month Professional Diploma : AI And Machine Learning Engineer Track

10/2023 – 07/2024

Information Technology Institute (ITI), MCIT - Alexandria

Bachelor's Degree In Data Science

09/2019 – 06/2023

Faculty of computers and data sciences, Alexandria University.

CGPA: 3.6797 /4 (Excellent)

TECHNICAL SKILLS

- **Programming language:** python, java, c++, R.
- **Deep Learning:** TensorFlow, Keras, PyTorch, spaCy, NLTK, surprise.
- **SQL:** Joins, sub-queries, views, indexes, triggers, procedures, Cursors, window functions.
- **Data Visualization and BI Tools:** Power BI, Tableau, Plotly-Dash.
- **Concepts:** NLP, Computer Vision, Recommender Systems, Generative AI, Transformers, LLMs, Time series.
- **Data Analysis, Classical ML, web scraping:** scikit-learn, Pandas, NumPy, Selenium, Streamlit.
- **Big Data & Stream Processing:** Hadoop, Spark (PySpark), Kafka, Hive.
- **Data Warehousing and ETL:** Talend.
- **Robotic Process Automation (RPA):** UiPath.
- **Tools & Technologies:** Cloud (AWS), Docker, Linux, Bash Scripting, MongoDB, Git, Agile (Trello).

PROJECTS

Blind Assistance System (Faculty Graduation Project) 🔗

- Developed a solution using smart glasses and a mobile app for blind individuals, incorporating object detection, text reading, scene description, cash/food recognition, and face recognition.
- **Research paper** :Smart Equipment for Assisting Blind People Published in (IJCSIS),Vol. 21, No.7 July 🔗
- Tools: **TensorFlow, pytorch, opencv.**

Comprehensive Fashion Recommender Systems (ITI / Twenty Too) 🔗

Developed three recommendation systems: similarity-based, personalization-based (appeal and look), and frequently bought together (complete the look) Using FP-Growth, CLIP, and Two-Tower Models, Vector Database and Deployed with Flask
Tools: mlxtend, Transformes, PyTorch, Pinecone, Flask

Head Pose Estimation (ITI) 🔗

- Predicted orientation angles (roll, pitch, and yaw) of human faces in images and videos using classical ML techniques.
- Tools: **OpenCV, MediaPipe, Scikit-learn, XGBoost, Streamlit.**

Necromancer: Colorization and Captioning for Videos and Images (ITI) 🔗

- Used DeOldify for artistic colorization of black-and-white media into vibrant images and videos. Implemented Salesforce's BLIP model for contextual captions and integrated audio using gTTS for enhanced multimedia experience
- Tools: **transforms, DeOldify, Opencv, torch, gTTS, Streamlit**

IMDb Analysis and recommendation system For Top Movies and Series (ITI) 🔗

- Developed end-to-end IMDb analysis pipeline, web scraping, data preprocessing, exploratory data analysis, dashboard development, recommendation system. **(Featured on Plotly Official Page)**
- Tools: **Selenium, Plotly Dash, Pandas, Scikit-learn**

Customer segmentation and anomaly detection (ITI) 🔗

- Conducted clustering and anomaly detection on a Credit Card Dataset, focusing on data preprocessing, dimensionality reduction, and data visualization, get business meaning, clusters description.
- Tools: **scikit-learn, scipy, pandas, matplotlib.**

Arabic Dialect Detection App (ITI) 🔗

- Implemented an Arabic Dialect Detection app to classify text into five distinct Arabic dialects using NLP techniques.
- Tools: **TensorFlow, sckit-learn, Streamlit, pandas**

CineMate: Movie Recommendation System app on MovieLens dataset (ITI) 🔗

- Implemented interactive user and item recommendation pages, utilizing Neural Collaborative Filtering for user recommendations and Item Similarity for movie recommendations.
- Tools: **TensorFlow, sckit-learn, Plotly-Dash, pandas**

Sentiment Analysis on Amazon Food Review Dataset (Classification and Deployment)

- Implemented sentiment classification on Amazon Food Review dataset and deployed the model.
- Tools: **pandas, scikit-learn, nltk, re, streamlit.**

Airlines Tweets Analysis Dashboard (ITI)

- Designed a dashboard to analyze customer sentiment towards airlines using Twitter data.
- Tools: **power bi.**

COURSES

- **Data Scientist Professional with Python (116 hours) - DataCamp**
- **Data Analyst in Power BI (48 hours) - DataCamp**
- **Data Analyst in SQL (39 hours) - DataCamp**
- **Data Analyst in Tableau (42 hours) – DataCamp**
- **Advanced Data Analysis - Udacity**
- **AWS Cloud Foundations - AWS Academy**

TRAINING

Trainee Data Scientist

07/2021 – 09/2021

Nawah.ai

- Applied machine learning/DL concepts such as feature engineering, modeling, and model evaluation & testing.
- Deployed machine learning projects as web apps using Streamlit.

LANGUAGES

- Arabic: Native
- English: Very Good