

Malak Soula

Egypt, Alexandria

☎ (+20) 1551021696 ✉ malaksola1606@gmail.com 🔗 [linkedin.com/in/MalakAbdelfattah](https://www.linkedin.com/in/MalakAbdelfattah)

🌐 <http://github.com/MalakSoula>

Education

university of Sadat City

Oct. 2021 – Jun. 2025

Faculty of computer science and Artificial intelligence, bioinformatics department

Sadat city, Egypt

Relevant Coursework

- | | | | |
|-----------------------|-----------------------|------------------------|-------------------|
| • Algorithms Analysis | • Genetic Algorithm | Systems | • Next-Generation |
| • Data analysis | • Biology | • Machine learning | Sequencing (NGS) |
| • Data structure | • Database Management | • Bioinformatics | Analysis |
| • OOP | | • Mathematical biology | |

Technical Skills

Languages: Python, SQL, R

Frameworks & Tools: Scikit-learn, SMOTE, TensorFlow, Seaborn, Pandas, NumPy, Power BI, Jupyter Notebook, Git, Galaxy, Mothur, Krona, Phinch, Google Colab, VS Code, Excel, GitHub, Matplotlib

Awards

SAS Curiosity Cup 2025 – Global Finalist (Top 15/112 Teams)

April 2025

- Finalist in SAS Curiosity Cup 2025 (Top 13% globally out of 112 teams); pitched healthcare analytics model using SAS tools.
- Presented project insights through a technical video pitch and storytelling format.
- Gained hands-on experience with SAS tools, forecasting, and real-world sustainability challenges.

SAS Forecasting Workshop and Hackathon for Sustainability and Data4Good

Dec. 2024

The Arab Academy for Management, Banking and Financial Sciences (AAMBFS)

Al Sheikh Zayed, Egypt

- Awarded 3rd place for developing a data-driven solution focused on Medical Inventory Optimization, addressing healthcare efficiency and sustainability.
- Collaborated with a team to solve real-world sustainability challenges using advanced forecasting techniques and data analysis.
- Recognized for innovation and practical application in a competitive environment supported by SAS and the ESG & Data4Good Center of Excellence.

Academic Excellence

Oct. 2022-2024

Faculty of computer science and arterial intelligence

Sadat City, Egypt

- Consistently achieved for academic and scientific excellence, achieving first place in the bioinformatics department with a GPA of 3.7 in my first year, maintaining a GPA of 3.5 in my second year, and achieving an impressive GPA of 3.67 in my third year. Awarded certificates for outstanding performance throughout these years.

Experience

Data Science and Machine Learning Trainee

Oct 2024 – Feb 2025

Microsoft Student Club - EELU

- Completed a 5-month intensive program covering Python, Pandas, Matplotlib, Power BI, and ML fundamentals.
- Built ML models including NASA NEO and Customer Segmentation projects.
- Designed and published interactive dashboards using Power BI, including:
 - 1- Pizza Sales Analysis - visualized category-, size- and time-based sales trends.
 - 2- Customer segmentation dashboard - combined ML output with business metrics for strategic targeting.
- Demonstrated real-world problem-solving with strong presentation and coding skills.

Cancer (Data Science) Internship

Sept. 2024 – Oct. 2024

HackBio

- Specialized in Data Science and Machine Learning for genomics, focusing on R programming and ML techniques.
- Analyzed genomic datasets and applied ML algorithms, documenting outcomes on GitHub.
- Enhanced computational biology skills through hands-on projects and teamwork.

Hamad medical cooperation

Aug. 2023 – Sep. 2023

Internship trainee

Doha, Qatar

- Played a pivotal role in the research process (systemic review research), contributing to a well-structured and comprehensive data foundation that earned co-authorship credit in the final publication
- Acquired practical hands-on experiences and engaging discussions with technical experts.
- Developed a profound understanding of the intersection between technology and healthcare, nurturing a passion for contributing to advancements in healthcare through technology integration.

Projects

Customer Segmentation Dashboard | *K-Means, Python, Scikit-learn, Pandas, Kaggle, Power BI*

Feb. 2025

- Segmented 5,000+ customers into 4 clusters using unsupervised learning; improved business targeting strategies and visualization using Power BI

NASA NEO Hazard Prediction | *Python, Scikit-learn, Pandas, NASA Dataset*

Jan. 2025

- Built NASA NEO hazard classifier with 94% precision on 338K+ records; improved detection of potential risks to Earth.
- Applied data processing, exploratory data analysis, and machine learning techniques to predict whether a NEO is hazardous or not.

Genetic Age Prediction | *Python, Kaggle, ML Models*

Jan. 2025

- Predicted biological age using gene expression data.
- Applied data processing, exploratory data analysis, and machine learning algorithms to predict age based on their genetic information

Vehicle Data Analysis (Training Project) | *Python, Kaggle*

Dec. 2024

- Analyzed vehicle pricing, mileage, and attributes using Python, with data sourced from Kaggle (Used Cars Dataset)
- Identified pricing trends for 10,000+ records and guided pricing strategy using Python-based EDA.
- Answered key business questions, including identifying top manufacturers/models and exploring price-mileage relationships.

Power BI Dashboard for Pizza Sales Analysis (Training Project) | *Power BI*

Dec. 2024

- Designed and implemented a Power BI dashboard using provided pizza sales data.
- Visualized key metrics, including sales by category, size, and date, to uncover trends and insights.
- Developed interactive features to enhance data exploration and reporting capabilities.

Cherry Tomato Classification – Gene Expression | *Python, kaggle, Bioinformatics*

Nov. 2024

- Classified organic vs non-organic cherry tomatoes using gene expression data
- Applied data preprocessing, exploratory data analysis, and ML algorithms for classification

Falcon's Football Club Management System | *Database Management, SQL Querying*

Dec. 2023

- Developed a software solution for Falcon's Football Club that streamlined club activities, including news updates, player and coach management, and team line-up evaluations.
- Utilized a sophisticated Database Management System (DBMS) to centralize information and enhance the user experience.
- Designed and implemented an Entity Relationship Diagram (ERD) to illustrate the relationships between different entities such as players, coaches, staff, and sponsors.
- Created a Relational Database Schema outlining the structure of tables, attributes, and relationships within the system.
- Created SQL diagrams and tables to represent the database structure and ensure data integrity and accuracy.

Sub-optimal health systematic review

Aug. 2023 – Sep. 2023

- Collaborated with the Hamad General Hospital surgical research team to develop a structured evidence base for a systematic review on sub-optimal health indicators. Contributed to literature review, data screening, and preliminary analysis.