

# Mohamed Haggag (Data Scientist)

---

[+201090414076](tel:+201090414076) | [mohamednasserabohamda@gmail.com](mailto:mohamednasserabohamda@gmail.com) | [GitHub](#) | [Portfolio](#)

<https://www.linkedin.com/in/mohamed-hagag-a117682a7> Egypt, Cairo

---

## Summary

### Computer Science & AI Student | Data Science & Machine Learning Specialist

Computer Science and Artificial Intelligence student with 3+ years of hands-on experience in advanced data analysis, business intelligence, and machine learning. Skilled in data cleaning, exploratory data analysis (EDA), and delivering actionable insights using SQL, Power BI, and Excel through advanced data visualization and reporting. Proficient in Python and experienced in designing and optimizing models using TensorFlow, Keras, and Scikit-learn, with strong expertise in neural networks (CNNs, RNNs) for solving real-world challenges. Experienced in natural language processing (NLP) for multilingual text analysis, and computer vision using YOLO and OpenCV for object detection, image segmentation, and OCR-based automation. Strong foundation in mathematics and statistics, including linear algebra, probability, and differential equations, for algorithm optimization. Familiar with cloud platforms (AWS, Azure, GCP), Docker, and API integration for real-time analytics. Additionally skilled in front-end development (ReactJS, HTML, CSS, JavaScript) for building interactive dashboards and deploying AI-driven web applications end-to-end.

---

## Skills

- **Programming:** Python (Pandas, NumPy, Scikit-learn), C++, JavaScript, SQL.
  - **Data Analysis & Visualization:** Power BI, Excel, Matplotlib, Seaborn, EDA, Statistical Modeling, Predictive Analytics
  - **Machine Learning & AI:** Predictive Modeling, Classification, Regression, Clustering, Dimensionality Reduction, TensorFlow, Keras, CNN, RNN, LSTM, Transformers.
  - **Computer Vision:** Image Classification, Object Detection, Image Segmentation, YOLO, OpenCV
  - **Natural Language Processing (NLP):** Text Preprocessing, Sentiment Analysis, Topic Modeling, Named Entity Recognition, Text Generation
  - **Big Data & Databases:** SQL, MongoDB, AWS, Azure, ETL.
  - **Web Development:** HTML5, CSS3, React.js.
  - **Tools & Platforms:** Git, GitHub, Jupyter Notebook, Docker, Kubernetes.
  - **Mathematics:** Linear Algebra, Calculus, Probability, Statistics.
  - **Analytical Skills:** Exploratory Data Analysis (EDA), Statistical Modeling, Predictive Analytics.
  - **Languages:** Arabic (Native), English (Intermediate).
  - **Soft Skills:** Problem-Solving & Analytical Thinking, Communication & Collaboration, Self-Learning & Adaptability, Efficiency & Organization, Work Ethic (Reliable, proactive, and committed to delivering high-quality results under deadlines)
- 

## Work Experiences

- **Machine Learning Trainee**  
**National Telecommunication Institute (NTI)**  
*Remote | 07/2025 - Present*
- **AI Engineer Freelance**  
**Dr Code**  
*Remote | 07/2025 - Present*
- **Data Analyst Intern**  
**Brainwave Matrix Solutions**  
*Remote | 02/2025 - 05/2025*
- **Data Analyst Intern**  
**Indolike**  
*Remote | 02/2025 - 05/2025*

---

## Education

Bachelor of Computers and Artificial Intelligence | Benha University – Egypt.

(Aug 2022 – expected Aug 2026)

---

## Projects

### Data Science & Machine Learning

- **Loan Approval Prediction** ([GitHub](#))  
Built an end-to-end ML pipeline (Python, Pandas, scikit-learn) for automated loan-eligibility scoring, from data cleaning to deployment via Streamlit. Improved model performance to **0.87 AUC** (↑ 17% from baseline 0.74) and reduced prediction latency by 25%.
- **E-Commerce Sales Prediction** ([GitHub](#))  
Developed a regression pipeline (XGBoost, CatBoost) to forecast sales, Achieved  **$R^2 = 0.90$** , improving forecast accuracy by 15% and enabling sales team to optimize discount strategy, increasing revenue by 8%. Engineered features like discount impact and deployed via Streamlit for real-time predictions.
- **Customer Churn Prediction** ([GitHub](#))  
Designed an ML pipeline (classification) to predict churn, covering EDA, feature engineering, and benchmarking. Predicted churn with 86% accuracy, enabling the retention team to reduce churn rate by 12% over 3 months. Deployed as an interactive Streamlit app for business use.
- **Parkinson's Disease Detection** ([GitHub](#))  
Built a classifier using biomedical voice data, following the full ML lifecycle. Deployed via Streamlit to enable real-time diagnostic support. Achieved 91% accuracy (F1 Score = 0.88) in detecting Parkinson's.
- **Bank Loan ML System** ([GitHub](#))  
Performed EDA and classification modeling (Jupyter Notebook) to predict loan approvals, Achieved 85% accuracy in predicting loan approvals and identified top 5 features influencing approval decisions, improving processing efficiency by 20%.

### Data Analysis & Visualization

- **Superstore Sales Analysis** ([GitHub](#))  
Cleaned data (Python), built a Power BI dashboard with DAX measures, and uncovered regional/category performance insights for growth strategies, Identified top-performing product categories leading to a 12% increase in targeted marketing ROI.
- **Pizza Restaurant Sales Dashboard** ([GitHub](#))  
Created an interactive Power BI dashboard to analyze sales trends, pricing impact, and customer preferences. Used Power Query for cleaning and DAX for KPIs. Insights led to a 10% increase in upselling through optimized pricing strategy.
- **Supermarket Sales Analysis** ([GitHub](#))  
Performed data cleaning, EDA, and integrated datasets (invoices, orders, sales team) to analyze sales trends, top products, and performance. Built interactive Power BI dashboards for actionable insights.
- **Car Theft Analysis in India** ([GitHub](#))  
Developed a Power BI dashboard to visualize theft patterns (geography, car models, recovery rates). Insights aided strategic security planning.
- **TikTok User Sentiment Analysis** ([GitHub](#))  
Analyzed app reviews (Python + Power BI) to extract sentiment trends, rating distributions, and common complaints. Processed 50K+ reviews, identifying top 3 complaint categories, leading to a 15% improvement in app store rating.

### Web Development

- **Portfolio Website** ([GitHub](#))  
Designed and deployed a responsive personal portfolio using HTML, CSS, and JavaScript.
- **E-Commerce Platform** ([GitHub](#))  
Built a front-end e-commerce platform with focus on UI/UX, leveraging HTML, CSS, and JavaScript

---

## Courses & Certification

- **Machine Learning, SQL, Power BI - DataCamp** ([Link](#))
- **2nd Place - GDG Banha Competition** ([Link](#))