

# Ahmed Ayman Al-Hofy

Data & ML Scientist | Cairo, Egypt | +201027557259 | ahmedalhofy42@gmail.com |  
<https://github.com/AhmedAyman4> | <https://www.linkedin.com/in/ahmed-alhofy/> | [View my website!](#)

## Summary

Aspiring Data & Machine learning Scientist with skills in Python, SQL, ML, NLP, and data analysis. Proficient in Pandas, Scikit-learn, Tableau, and Power BI. Information Systems student at the Faculty of Computers & AI, University of Sadat City, Passionate about using data to solve everyday challenges and applying AI to create practical, impactful solutions.

## Education

Bachelor's Degree, Faculty of Computers & Artificial Intelligence – University of Sadat City

Graduating June 2026

CGPA: 3.23 & Major: Information Systems.

## Skills

- Python (NumPy | Pandas | Scikit-learn | TensorFlow | NLTK | LangChain | LangGraph) | SQL | Shell Scripting (Bash) | HTML | CSS | Docker | Git | GitHub | Excel | Tableau | Power BI
- Machine Learning | NLP | LLM Applications | Prompt Engineering | Data Analysis | Data Visualization | Statistics | Probability | OOP | Communication | Problem-Solving | Adaptability | Analytical Thinking

## Experience

### AI Engineer | Konecta · Internship

New Cairo, Egypt | Jul 2025 - Present

- Conducted data collection and preprocessing using Python libraries and web scraping tools to prepare structured datasets for analysis.
- Built and optimized machine learning models for classification, clustering, and image recognition with tuning and evaluation.
- Built RAG-based chatbots with document indexing, embeddings, and multi-turn query handling.
- Built automation workflows using n8n to streamline data and AI processes.

### Data scientist | Digital Egypt Pioneers Initiative - DEPI · Internship

Alexandria, Egypt | Oct 2024 – May 2025

- Analyzed data with Python, SQL, and analytics tools to extract insights.
- Built and deployed machine learning models using Scikit-learn, MLflow, and prompt engineering.

### Intern | CIB Egypt · Internship

Cairo, Egypt | Jul 2024 - Jul 2024

- Gained data and financial literacy expertise, entrepreneurship, and teamwork to enhance problem-solving, decision-making, and business success.

## Projects

### Semantic Book Recommender With LLMS:

- Built a semantic book recommender with natural language search, vector embeddings, and sentiment analysis for theme- and emotion-based discovery.
- Implemented zero-shot classification to tag 7,000+ books as fiction or non-fiction, enabling filterable, precise searches.
- Developed a Gradio web app to integrate NLP pipelines into an interactive dashboard, cutting manual dataset exploration by 80%+.

### FCAI Regulations Chatbot:

- Built a bilingual RAG chatbot with LangChain, Gemini, and Gradio to deliver Arabic-English, source-based answers from FCAI regulations.
- Optimized retrieval and embeddings using ChromaDB and HuggingFace, improving response speed and accuracy.

### Customer Segmentation:

- Clustered wholesale customers using K-Means, Hierarchical, DBSCAN, and MeanShift to support targeted strategies.
- Preprocessed and visualized data with scaling, log transforms, and PCA; evaluated clusters using multiple indices.
- Derived customer profiles from spending patterns to guide marketing and sales insights.

### Automated Market Data Scraper (Real Estate & E-Commerce):

- Developed Python scrapers (Playwright, BeautifulSoup) to extract 1,400+ real estate listings into structured Pandas DataFrames.
- Automated infinite scrolling and pagination, reducing manual effort by over 90%, and delivered versioned CSVs with UTF-8 encoding for analysis.

### Invoice Reader App:

- Built an AI-powered invoice reader using Tesseract OCR, Google Gemini, and LangChain, automating data extraction and structuring from invoices, which reduced manual entry time by 80%.
- Developed Flask and Streamlit web apps enabling real-time OCR processing and Gemini-based text-to-JSON conversion, which improved data accuracy and enhanced user experience for non-technical users.
- Integrated MongoDB to store extracted data with embeddings for vector search and connected it to a chatbot for intelligent querying of invoices and documents.

Courses And Certifications

---

• SAP Certified Associate - SAP Generative AI Developer Issued by SAP	Sep 2025
• AWS Certified AI Practitioner Issued by Amazon Web Services Training and Certification	Apr 2025
• Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate Issued by Oracle	Jun 2025
• Data Scientist Associate Issued by Data Camp	Aug 2024
• IBM SkillsBuild Data Analytics Certificate Issued by IBM SkillsBuild	Jun 2025
• AWS Academy Graduate - AWS Academy Machine Learning for Natural Language Processing	Aug 2024
• AWS Academy Graduate - AWS Academy Cloud Foundations	Jun 2024
• AWS Academy Graduate - AWS Academy Data Engineering	Jun 2024
• Applied Data Science with Python - Level 2 Issued By IBM	May 2024
• Google Data Analytics Professional Certificate Issued by Coursera	Jul 2024
• The Machine Learning Process A-Z – Issued By 365 Data Science	Nov 2024
• McKinsey Forward Program - Issued by McKinsey & Company	Dec 2024