

# Ahmed Noshy

Gen Ai Engineer, Data Scientist

Cairo, EG | Phone: +201063944671 | [Ahmed.noshy2004@gmail.com](mailto:Ahmed.noshy2004@gmail.com)  
[Github.com/NASO7Y](https://github.com/NASO7Y) | [linkedin.com/in/nos7y](https://linkedin.com/in/nos7y)

## SUMMARY

AI Engineer specializing in Generative AI, Data science, and intelligent automation — with hands-on experience building production-grade LLM applications, computer vision systems, and end-to-end ML pipelines. Proven ability to deliver high-accuracy models and scalable AI workflows, NLP, and enterprise automation domains.

## EDUCATION

### Bachelor of Artificial Intelligence

MENOUFIA UNIVERSITY

Faculty of Artificial Intelligence, Machine Intelligence Department.

Grade: B

Menoufia, Egypt

October 2022 - July 2026

## LANGUAGE

Arabic (Native), English (C1)

## WORK EXPERIENCE

### Generative Ai & Automation Engineer

[Hatif](#) - Remote - Riyadh, SA

June 2025 - Present

- **Architecting and enhancing the advanced voice system** – leading the development and refinement of conversational AI systems to ensure high accuracy, natural interaction - human-like-, and seamless user experience.
- **Developing enterprise-grade AI solutions**– designing and deploying tailored intelligent systems that drive automation, streamline operational workflows, and eliminate redundant manual tasks.
- **Building intelligent AI-powered chatbots for enterprise clients** – leveraging LLMs to automate customer interactions and deliver personalized, context-aware responses at scale.

### AI Research Assistant

Independent - Remote - Cairo, EG

March 2025 - May 2025

- **Conducting Master's level research on predictive healthcare** – Developed a heart disease prediction system using ensemble ML models (Random Forest, XGBoost, SVM), achieving 95% accuracy and 97% recall on patient records, outperforming baseline clinical screening benchmarks.
- **Building an interactive diagnostic tool** – integrating Random Forest, XGBoost, SVM, and other optimized models into a user-friendly Streamlit UI with cross-validation and hyperparameter tuning for robust real-world deployment.

### Independent Data Scientist

Freelancer - Remote - Cairo, EG

Jan 2025 - Present

- **Collaborating with clients on automation solutions** – creating enterprise-ready AI agents and workflow optimizations that enhance efficiency and customer engagement.
- **End-to-end AI project delivery** – handling the full cycle from data collection and preprocessing to model training, evaluation, and deployment on interactive platforms (Streamlit, FastAPI).
- **Developing AI-powered computer vision systems** – delivering solutions such as disease detection in medical images, and anti-cheating face and motion detection with high accuracy and reliability

## TECHNICAL SKILLS

---

**Programming Languages:** Python, C++

**Data Exploration:** Pandas, Numpy, matplotlib, Seaborn, Plotly

**ML/DL:** Scikit-learn, TensorFlow, PyTorch, Keras, Hugging Face, XGBoost, LightGBM

**CV:** OpenCV, YOLO

**NLP:** spaCy, NLTK, Transformers

**Automation:** n8n, Langchain, Make

**Model Deployment:** Streamlit, FastAPI, PyQt, Docker

## SOFT SKILLS

---

Problem-Solving, Critical Thinking, Teamwork, Adaptability, Time Management, Leadership

## PROJECTS

---

### EGYPTIAN ID OCR

- Built a production-grade OCR pipeline for Egyptian National ID cards, achieving 90% Arabic text extraction accuracy with integrated fraud detection and identifying manipulated images and invalid ID patterns using YOLOv11 and custom-trained models.
- **Tech Stack:** Python, OpenCV, YOLOv11, AI model training, OCR model, Streamlit  
**Link:** [github.com/NASO7Y/OCR\\_Egyptian\\_ID](https://github.com/NASO7Y/OCR_Egyptian_ID)

### 10+ AUTOMATION PROJECTS

- A collection of automation tools and AI solutions designed to streamline repetitive tasks, improve productivity, and integrate smart data processing workflows. Projects include document automation, voice-based assistants, dashboard generators, and more.
- **Tech Stack:** n8n, Make, Zapier, JavaScript, HTML, LLMs.  
**Link:** [github.com/NASO7Y/My-Automation-Projects](https://github.com/NASO7Y/My-Automation-Projects)

### ATS (Applicant Tracking System)

- Implemented an Applicant Tracking System that analyzes resumes using NLP + Computer Vision. Features include resume parsing (OCR for PDFs/images), job category prediction across 25+ classes, keyword matching, section detection, plus a FastAPI and Streamlit GUI frontend.
- **Tech Stack:** Python, FastAPI, Streamlit, OpenCV, YOLO, OCR (PyMuPDF + Vision Agent), NLTK, AI Model Training  
**Link:** [github.com/NASO7Y/ATS--NLP--project](https://github.com/NASO7Y/ATS--NLP--project)

### DEVELOPING A RAG SYSTEM

- Built a RAG system using DeepSeek R1, Qwen 2.5, and Ollama, optimizing retrieval pipelines for ~85% accuracy and context-aware responses.
- **Tech Stack:** Python, OpenCV, YOLOv11, AI model training, OCR model, Streamlit  
**Link:** [github.com/NASO7Y/RAG-PDF-Search](https://github.com/NASO7Y/RAG-PDF-Search)

### CHEST X-RAY DETECTION

- Developed and deployed a deep learning model for automated disease detection in chest X-ray images, achieving 93% accuracy. Utilized advanced CNN architectures to improve diagnostic precision and implemented real-time inference for medical use cases.
- **Tech Stack:** Python, YOLOv11, MobileNet, ResNet50, EfficientNet, PyTorch, Streamlit, FastAPI, Deployment  
**Link:** [github.com/NASO7Y/Chest-X-ray-Detection](https://github.com/NASO7Y/Chest-X-ray-Detection)

### REAL DATA SENTIMENT ANALYSIS

- Developed a Twitter sentiment analysis system to classify tweets as positive, negative, or neutral, achieving 91% accuracy. Applied NLP and ML techniques to track public opinion trends and generate insights for social media monitoring and brand analysis.

- **Tech Stack:** Python, PyTorch, Streamlit, FastAPI, Deployment  
**Link:** [github.com/NASO7Y/Twitter-Sentiment-Analysis](https://github.com/NASO7Y/Twitter-Sentiment-Analysis)

## STUDENT PERFORMANCE ANALYSIS

- Developed a machine learning-based system to analyze student performance and predict outcomes. Created an interactive Power BI dashboard for real-time insights and progress tracking.
- **Tech Stack:** Python, Pandas, Scikit-learn, Matplotlib, NumPy, Seaborn, SQL, Excel  
**Link:** [github.com/NASO7Y/Students-Performance-Analysis](https://github.com/NASO7Y/Students-Performance-Analysis)

## CERTIFICATES

---

### **Kaggle: Python**

Core Topics: Python syntax, functions, booleans, conditions, loops, lists, strings, dictionaries, working with external libraries

Certificate: [kaggle.com/learn/certification/nos7yy/python](https://kaggle.com/learn/certification/nos7yy/python)

### **Kaggle: Pandas**

Core Topics: Reading/Writing, Indexing & Assigning, Summary functions, Maps, Grouping and Sorting, Data types, Missing Values

Certificate: [kaggle.com/learn/certification/nos7yy/pandas](https://kaggle.com/learn/certification/nos7yy/pandas)

### **Kaggle: Introduction to Machine Learning**

Core Topics: How models work, Basic Data Exploration, Model Validation, Underfitting and Overfitting, ensemble learning

Certificate: [kaggle.com/learn/certification/nos7yy/intro-to-machine-learning](https://kaggle.com/learn/certification/nos7yy/intro-to-machine-learning)

### **Datacamp: Supervised Learning with scikit-learn**

Core Topics: Classification, Regression, Fine-Tuning, Preprocessing, and Pipeline,

Skills: Scikit-learn, Cross-validation, GridSearchCV, Decision Trees, SVMs, Random Forest

Certificate: [datacamp.com/completed/statement-of-accomplishment/course/ebc8ff7baa1e2870e7c1b339921f31e7c077edca](https://datacamp.com/completed/statement-of-accomplishment/course/ebc8ff7baa1e2870e7c1b339921f31e7c077edca)

### **Machine Learning Specialization: Stanford University/DeepLearning.AI**

Core Topics: Supervised Machine Learning, Unsupervised machine learning, Advanced learning algorithms

Certificate: <https://coursera.org/share/a9d85343d2d3a3bd39b2f5b857adf686>

### **Google: Introduction to Large Language Models**

Core Topics: Transformer Architecture, Pretraining & Fine-tuning, Prompt Engineering, Model Optimization

Certificate: [coursera.org/share/0c3e0a97adc8272a5ceb3d257e760e3d](https://coursera.org/share/0c3e0a97adc8272a5ceb3d257e760e3d)

### **Google: Introduction to Generative AI**

Core Topics: GenAI as a subset of DL, Hallucinations, Prompts, Google Cloud for Enhanced GenAI

Certificate: [coursera.org/share/42c059fad09aeb3b30943f7031697599](https://coursera.org/share/42c059fad09aeb3b30943f7031697599)

## ACTIVITIES

---

### **Supervisor in Problem Solving Track**

**IEEE Menoufia Student Branch - Hybrid,**

**Menoufia, EG**

- Designed and implemented problem-solving sheets, overseeing project development
- Ensured participants gain valuable experience.
- Troubleshoot and resolved technical issues