Moustafa Mohamed

AI Engineer | Specializing in Machine Learning, Deep Learning & LLM Engineering
moustafa.mh.mohamed@gmail.com| +90 536 654 09 89

Linkedin | Github | Portfolio | Kaggle

Istanbul, Türkiye

About Me

I am a second-year Software Engineering student with a strong foundation in programming, data analysis, and machine learning. Proficient in Python, C++, and SQL, I have hands-on experience utilizing data visualization libraries such as Matplotlib and Seaborn. My technical expertise includes designing, building, and deploying machine learning and deep learning models across diverse, data-driven applications.

I have earned certifications from recognized institutions, including IBM's *Python for Data Science, AI & Development*, Udemy's *Python for Machine Learning & Data Science Masterclass*, and *Deep Learning A-Z 2025: Neural Networks, AI & ChatGPT Prize*. Currently, I am advancing my knowledge in Generative AI and Large Language Models (LLMs) through the *LLM Engineering Masterclass* and Google Cloud's *Generative AI Learning Paths*.

Driven by a strong passion for artificial intelligence, natural language processing, and intelligent automation, I am dedicated to creating innovative and impactful AI solutions. My ambition is to contribute meaningfully to the forefront of AI research and development, focusing on LLMs and real-world NLP applications that transform how we interact with technology.

Education

Istanbul Topkapi University — Bachelor of Science in Software Engineering
Istanbul, Türkiye | Sep 2023 – Present

Istanbul Topkapi University — Turkish Language and Literature
Istanbul, Türkiye | Feb 2023 – Sep 2023

Skills

Programming Languages:
Python, C, C++, JavaScript, SQL

Data Analysis, Machine Learning & Deep Learning:

- **Libraries & Tools:** NumPy, Pandas, Matplotlib, Seaborn, Plotly, SciPy, scikit-learn
- **Core Competencies:** Exploratory Data Analysis (EDA), Data Cleaning, Data Manipulation, Feature Engineering
- **Machine Learning:** Supervised & Unsupervised Learning, Model Evaluation, Clustering, Regression, Classification

- **Deep Learning:** Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN)
- **Frameworks:** TensorFlow, PyTorch, Keras

LLM & Generative AI:

- **Core Concepts**: Prompt Engineering, Tokenization, Retrieval-Augmented Generation (RAG)
- **Tools & Frameworks**: LLaMA, Google Gemini API, OpenAI API

Development & Engineering Tools:
Git, GitHub, Jupyter Notebook, Power BI, Visual Studio Code, PyPI (Python Packaging)

Programming Concepts & Practices:
Object-Oriented Programming (OOP), Software Packaging, API Integration, NLP (Natural Language Processing)

Projects & Libraries:

- Creator of **datacmp** – a Python library for data cleaning and EDA published on PyPI
- Built AI projects including spam classifiers, image recognition with CNNs, clustering models, and summarization tools using LLMs

Projects

- **SMS Spam Classifier**
  Developed a machine learning model to classify SMS messages as spam or ham using text preprocessing, feature extraction (TF-IDF), and logistic regression.
  [GitHub Repository]
- **Fruits & Vegetables Image Recognition**
  Built a deep learning model using Convolutional Neural Networks (CNN) to classify images of fruits and vegetables. Utilized TensorFlow and Keras for model training and evaluation.
  [GitHub Repository]
- **Mall Customer Segmentation**
  Applied unsupervised learning (K-Means Clustering) to segment mall customers based on behavioral data for targeted marketing strategies.
  [GitHub Repository]
- **San Francisco Salaries Data Analysis**
  Conducted exploratory data analysis (EDA) on public employee salary data, uncovering insights through data cleaning, visualization, and statistical analysis.
  [GitHub Repository]
- **Titanic Dataset Analysis**
  Performed end-to-end data analysis including EDA, hypothesis testing, and regression modeling to predict passenger survival.
  [GitHub Repository]

- **Web Summarizer AI**
  Developed an intelligent summarization tool using Python, Selenium, BeautifulSoup, Google Gemini API, and LLaMA 3.2 to extract and summarize web content dynamically.
  [GitHub Repository]

- **datacmp – Python Library for EDA & Data Cleaning**
  Created and published a lightweight Python library for automated exploratory data analysis and data preprocessing using pandas. Available on PyPI.
  [GitHub Repository] [PyPI]

- **Car Price Prediction**
  Car Price Prediction using Linear Regression. A machine learning model that predicts car prices based on features like name, company, year, kilometers driven, and fuel type. Built with Linear Regression and deployed as a Streamlit web app.
  [GitHub Repository]

- **Machine Learning Projects Portfolio**
  A consolidated repository of machine learning projects demonstrating skills in classification, regression, clustering, and visualization using real-world datasets and tools.
  [ML Projects Repository]

Certifications

- **IBM's "Python for Data Science, AI & Development"** – Coursera
- **Python for Machine Learning & Data Science Masterclass** – Udemy
- **Deep Learning A-Z 2025: Neural Networks, AI & ChatGPT Prize** – Udemy
- **AI Python for Beginners** – DeepLearning.AI
- **Deep Learning MiniCamp [Arabic]** - Udemy
- **LLM Engineering: Master AI, Large Language Models & Agents** – Udemy
- **Beginner: Introduction to Generative AI Learning Path** – Google Cloud (In Progress)

Languages

- **Arabic -** Native
- **English -** Advanced (Professional working proficiency)
- **Turkish -** Intermediate (Conversational)