MUHAMMAD GHOUS

House no: C-707 Majeed Colony Sector 1, Landhi, Karachi , +923002514961, mghous373@gmail.com
https://www.linkedin.com/in/muhammad-ghous-01ab29255/

Profile:

Passionate Artificial Intelligence graduate ready to apply comprehensive knowledge in Machine Learning, Deep Learning, NLP, and Generative AI to real-world challenges. Adept in Python and its ecosystem, I'm seeking a role to contribute to innovative AI projects and develop practical skills in a dynamic professional environment.

Education:

Dawood University of Engineering & Technology **(NOV 2021 – 2025)**
Bachelor of Science in Artificial Intelligence, CGPA 3.34 out of 4.00

Skills:

i. Python
ii. Machine Learning
iii. Google Gen AI
iv. Data Analysis
v. Deep Learning
vi. PyTorch
vii. Data Visualization
viii. NLP
ix. SQL
x. Fast API
xi. Hugging Face
xii. Computer Vision

Certification & Achievements:

i: Complete a course on Google IT Automation with Python from Google Coursera.
ii: Complete a course on Google Data Analytics from Google Coursera.
iii: Complete a Data Scientist Associate certificate from Data Camp.
iv: Complete a course on Google Prompting Essentials from Google Coursera.

Internships and Training:

- Completed a Data Science remote internship at Digital Empowerment Network (July 2025 to Sep 2025).
- Completed a Data Science remote internship at AA Solutions (August 2025).
- Worked as a Trainee AI Engineer at Wenawa for 2 months (Oct 2025- Dec 2025)

Projects:

- **Interactive Driven AI Learning Platform (Final Year Project)**

The final year project is in the process. While considering SDGs point, we aimed to build a Virtual AI Mentor which works on the four base models to receive and answer the query of the user, four models are: Text-to-Text, Speech-to-Text, Text-to-Speech and Speech-to-Speech. By the help of these models the program will be able to take input in the form of text or speech and give output in the form of text or speech. In this project, we avoid to use already built API's instead we are focusing on the transformers and pre-trained models.

- **Hate Speech Detection:**

Built an NLP model to classify social media comments as hate or non-hate using Count Vectorizer, TF-IDF, and Logistic Regression. Achieved high accuracy and F1 score on a real world dataset after preprocessing and vectorization.

- **Sentiment Analysis on IMDB Dataset:**

Developed a sentiment classifier using NLP and ML techniques to identify positive or negative reviews. Pre processed raw text, extracted features using TFIDF, and trained classifiers like Logistic Regression and SVM with high accuracy.

- **SpaceX First Stage Landing Prediction:**

Developed a SpaceX First Stage Landing Prediction system using Python, Machine Learning (Random Forest with GridSearchCV), and Streamlit. Integrated live data from the SpaceX API, performed exploratory data analysis with Plotly & Folium, and built an interactive dashboard to predict landing success based on payload, orbit, rocket type, and launchpad.

- **Email Spam Classifier:**

Developed a machine learning model to classify emails as spam or ham using Natural Language Processing (NLP) techniques. Preprocessed text data by removing stopwords, punctuation, and applying stemming. Extracted features using TF-IDF Vectorizer and trained a Multinomial Naive Bayes classifier. Evaluated model performance using accuracy, precision, and recall metrics. Deployed the project using Streamlit for an interactive web-based interface.

- **House Price Predictor:**

Developed a supervised machine learning model to predict house prices using a real-world dataset. Performed extensive data preprocessing including handling missing values, encoding categorical variables, and creating derived features (e.g., house age, price per square foot). Conducted exploratory data analysis and outlier detection to understand key pricing drivers. Trained and

evaluated multiple regression models (Linear Regression, Random Forest, Gradient Boosting) and selected the best-performing model using R^2 score and MSE. Deployed future price prediction functionality using user-defined property features.

- **Karachi Air Quality and Health Prediction System:**

Designed and deployed an end-to-end Karachi Air Quality & Health Prediction System integrating Flask RESTful APIs, XGBoost (Bayesian optimization), and real-time data pipelines. Built an interactive Streamlit dashboard with predictive analytics, health risk stratification, and environmental trend visualization. Implemented microservices architecture, Docker deployment, and CI/CD for scalable, production-grade performance.

- **WhatsApp Chat Analyzer:**

Created an interactive web app using Streamlit to analyze WhatsApp chat data, including most active users, emoji statistics, and message trends. Implemented 24-hour time format and generated dynamic visualizations using matplotlib and seaborn.

- **PDF Book Reader with AI Chat:**

Developed a PDF Book Reader with AI Chat using Streamlit and Python, allowing users to upload PDFs and interact with them through an AI-powered chat interface. The application leverages RAG (Retrieval-Augmented Generation) with ChromaDB for vector storage and Sentence Transformers for embeddings, while integrating Google Gemini API to generate accurate, context-aware responses. A clean and interactive Streamlit UI was built to support real-time chat and document processing, supported by a pipeline for PDF text extraction, chunking, embedding, similarity search, and AI response generation. Additionally, robust error handling was implemented for API, PDF, and database issues, with planned enhancements including multi-PDF support, summarization, and exportable chat history.

- **WhatsApp Food Ordering Chatbot:**

Developed an AI-powered WhatsApp Food Ordering Chatbot using FastAPI and Python that enables customers to place orders via text, voice, and interactive buttons. Integrated NLP and speech-to-text models (HuggingFace Wav2Vec2) for accurate order understanding in English and Urdu, and implemented location-based routing using OpenStreetMap geocoding with the Haversine formula to identify the nearest restaurant branch. Designed a Streamlit admin dashboard for real-time order management, menu control, and branch monitoring while building a complete end-to-end system including database architecture, webhook handling, conversation tracking, and real-time WhatsApp Business integration via GREEN-API.

- **OCR Scan Project:**

Developed an advanced OCR system using PaddleSOCR and deep learning models for high-accuracy text extraction from scanned documents and medical reports, implementing noise reduction, confidence scoring, and multi-model comparison to measure accuracy, speed, and reliability. Integrated BERT-based NLP pipelines for medical entity recognition, semantic search, and intelligent text understanding, while building a FastAPI backend with endpoints for single/batch OCR processing, medical search, and entity statistics. Designed a complete analytics workflow to generate performance metrics, entity summaries, and comparison visualizations for real-time evaluation and validation.