

# Muhammad Usman Haider

## Data Scientist

• Lahore, PK • +92-3041581788 • [osmanhaider167@gmail.com](mailto:osmanhaider167@gmail.com) • [LinkedIn](#) • [Github](#)

### Professional Summary

Experienced **ML Engineer** with **1.9 years of experience**, specializing in Generative AI LLMs. Proficient in developing and deploying cutting-edge AI solutions within SaaS and PaaS environments.

### Professional Experience

#### STECHAI

Sep 2023 - Present

A service-based company delivering comprehensive data science, AI, and machine learning solutions.

##### *Associate Data Scientist – Onsite*

- Engineered high-performance Django and FastAPI backend services, increasing user traffic by 50%.
- Built and fine-tuned open-source ML and deep learning models (NLP, CV), achieving 15% higher predictive accuracy and automating 65% of repetitive tasks.
- Developed a Retrieval-Augmented Generation (RAG) chatbot using LangChain, Pinecone, and GPT-based models to provide precise legal assistance.
- Deployed the solution using Docker and AWS, ensuring high availability, scalability, and efficient resource utilization in a production environment.

### Projects

**Student Medical Chatbot** – Developed a chatbot app for MBBS students to assist with their medical books using Llama Index. Improved accuracy by fine-tuning the model and storing embeddings in OpenSearch on AWS Cloud. Built the backend in Django with proper authentication and session management.

**Chatical** - I led a project where we built a Chrome extension that automatically captures the content of the current webpage when you open it. The content is then stored in a Pinecone vector database, allowing users to chat and ask questions about the article they're viewing. The backend was developed using FastAPI and used a Retrieval-Augmented Generation (RAG) approach for answering questions.

**TrainMyAI** - I develop a user-friendly tool designed to enable anyone to train deep learning models with just a single click. Built with a simple and interactive interface powered by Streamlit, this project aims to simplify the process of model training, evaluation, and inference.

**Inventory Analysis** – Led the project and trained an XGBoost model on a dataset with over 5 million rows. Developed Django APIs to streamline model training and prediction for both sales forecasting and replenishment optimization. Integrated with AWS for data storage and retrieval.

**Football Analysis** – Developed a football analysis system that identifies player positions and tracks their movement speed in real time during live matches. We used YOLO for object detection, OpenCV for video processing, and PyTorch for building the underlying models.

### Technical Skills

- **Programming & Development:** Proficient in Python, Scikit-Learn, TensorFlow, Keras, PyTorch, NLTK, Transformers, EDA, ETL, Matplotlib, Plotly, OpenCV, SQL, Vector Databases, Flask, Django Streamlit, Docker, GitHub Actions, CI/CD pipelines, NoSQL, DSA.
- **Data Science & AI:** Expertise in Machine Learning, Deep Learning, Data Analysis, Predictive Modeling, Artificial Intelligence, Statistical Modeling, Data Visualization, Data Engineering, Model Deployment, Time Series Analysis, Cloud Machine Learning, Big Data Technologies, Natural Language Processing (NLP), Prompt Engineering, Generative AI, LangChain, Llama-Index, Retrieval-Augmented Generation (RAG) with both open source and paid LLM models.
- **Cloud Services:** AWS Cloud, Microsoft Azure, Google Cloud
- **Communication & Collaboration:** Strong in Effective Communication, Collaboration and Team Building.
- **Languages:** Fluent in English and Urdu with full professional proficiency.

### Certifications

#### Multi AI Agents and Advanced Use Cases with crewAI – Deep Learning

- [Multi AI Agents and Advanced Use Cases with crewAI](#)