

Muhammad Mohsin

Data Scientist & Machine Learning Engineer

Faisalabad, Pakistan | mmohsin4901@gmail.com | +92 325 6111094

github.com/Mohsin-Irfan | linkedin.com/in/mohsinirfan1

PROFESSIONAL SUMMARY

Data Scientist with 4+ years of progressive experience delivering ML-driven solutions across predictive analytics, NLP, computer vision, and business intelligence. Proven track record of improving model accuracy by up to 35%, reducing reporting overhead by 50%, and driving measurable revenue growth. Experienced building scalable pipelines on Google Cloud, with a strong foundation in stakeholder communication and cross-functional delivery.

TECHNICAL SKILLS

Languages & Databases: Python, SQL, MongoDB

ML & Analytics: Scikit-learn, TensorFlow, OpenCV, NLP, Predictive Modeling, Recommendation Systems, Clustering, A/B Testing, Revenue Forecasting

Data Visualization: Power BI, Tableau, Matplotlib, Seaborn, Plotly

Libraries & Frameworks: NumPy, Pandas, Flask, Streamlit

Cloud & DevOps: Google Cloud AI Platform, Git, API Integration

PROFESSIONAL EXPERIENCE

Hive Technologies | Data Scientist

Jul 2024 – Present

- Architected end-to-end predictive analytics pipelines that improved cross-departmental decision-making speed by 30%.
- Designed and deployed interactive Power BI dashboards adopted company-wide, eliminating 50% of manual reporting effort.
- Refined time-series revenue forecasting models, lifting forecast accuracy by 20% and enabling more reliable quarterly planning.

Turing | Senior Data Scientist (Contract)

Sep 2024 – Nov 2024

- Re-engineered production ML models through targeted feature engineering and hyperparameter tuning, achieving a 35% accuracy improvement.
- Deployed AI automation workflows that increased overall system efficiency by 40%, reducing processing latency for downstream analytics.
- Delivered scalable model infrastructure on Google Cloud AI Platform, supporting high-availability production workloads.

Hive Technologies | Senior Data Analyst

Jun 2022 – Jul 2024

- Built ML-based customer segmentation models that enabled targeted marketing campaigns and improved customer retention strategies.
- Overhauled BI dashboards to support multi-department use cases, consolidating reporting across 3+ business units into unified views.
- Automated recurring analytics workflows, significantly reducing analyst hours and improving data pipeline reliability.

Hive Technologies | Junior Data Analyst

Apr 2021 – Jun 2022

- Automated SQL-based reporting workflows, reclaiming 15+ hours per week previously spent on manual data pulls.
- Developed predictive sales forecasting models that contributed directly to 12% revenue growth over the following fiscal period.
- Strengthened data pipeline integrity and collaborated cross-functionally to align analytics outputs with business requirements.

KEY PROJECTS

MirrorWorld — Real-Time AI Gesture & Expression Engine

Built a real-time AI system that transforms a standard webcam into an interactive experience: hand gestures summon visual effects (fire, lightning, ice) and facial expressions dynamically alter the environment. Demonstrates applied computer vision at the intersection of ML and creative interfaces.

Tech: Python, OpenCV, MediaPipe, Real-Time Computer Vision, Facial Expression Recognition

CineVerse — AI-Powered Movie Recommendation System

Engineered a content-based recommendation engine using TF-IDF vectorization and cosine similarity to deliver personalized movie suggestions. Integrated the TMDB API for real-time poster, rating, and trailer data; deployed as an interactive Streamlit web app.

Tech: Python, Streamlit, Scikit-learn, NLP, TF-IDF, Cosine Similarity, TMDB API

Hand Gesture Calculator

Developed a touchless calculator controlled entirely through hand gestures captured via webcam. Leveraged real-time hand landmark detection to map finger positions to arithmetic operations — no keyboard or mouse required.

Tech: Python, OpenCV, MediaPipe, Computer Vision

DataWiz Pro — End-to-End Analytics Application

Designed a self-service analytics platform enabling non-technical users to upload datasets (CSV/Excel), clean and preprocess data with intelligent automation, build interactive visualizations, and run built-in ML models for outcome prediction.

Tech: Python, Streamlit, Pandas, Scikit-learn, HTML/CSS/JavaScript, API Integration

Fake News Detector — NLP Classification Model

Trained a misinformation classification model achieving 92% accuracy on a benchmark dataset. Applied TF-IDF feature extraction and ensemble classifiers, with EDA visualizations to surface patterns in real vs. fake news corpora.

Tech: Python, Pandas, Scikit-learn, NLP, TF-IDF, EDA

Face Attendance — Biometric Attendance System

Built a real-time facial recognition system for automated attendance tracking, replacing manual entry with high-accuracy biometric identification. Supports automatic logging and reporting for organizations and educational institutions.

Tech: Python, OpenCV, Facial Recognition

EDUCATION

BS Cyber Security | Riphah International University, Pakistan | 2025 – 2029 (In Progress)

Intermediate — ICS | Superior College, Pakistan | Completed Aug 2025

CERTIFICATIONS

- **Machine Learning Specialization** — DeepLearning.AI (2025)
- **Power BI Data Analyst** — Microsoft (2024)
- **Introduction to Statistics** — Stanford University (Online) (2023)
- **Scientific Computing with Python** — freeCodeCamp (2022)

LANGUAGES

English — Fluent **Urdu** — Native