Muhammad Usman

Saarbrücken, Germany | +49 174 276 7793 | usmanchaudhary.eng@gmail.com |
linkedin.com/in/usmanbioinformatician2025 | github.com/Muhammad-Usman678

Professional Summary

Bioinformatician and **DevOps-oriented Machine Learning Engineer** with experience delivering scalable solutions across **genomics, transcriptomics, and proteomics**. Skilled in **reproducible NGS/ML pipeline development**, workflow automation, and CI/CD-driven deployment across **Linux and HPC environments**. Strong background in **Python, R, and C++**, with expertise in **multi-omics data integration, ETL workflows, knowledge graphs (Neo4j)**, and developing dashboards and data products for research and clinical decision support.

Skills

NGS & Bioinformatics: WGS, WES, RNA-seq (bulk & single-cell), methylome, variant calling, annotation, Bioconductor
ML & Data Science: Scikit-learn, PyTorch, PCA/UMAP/t-SNE, clustering, Random Forest, SVM, XGBoost, SHAP, GNNExplainer, graph ML
DevOps & Infrastructure: Docker, Git, Bash, Linux Server Management, CI/CD, Workflow Automation, Cloud Deployment
Pipelines & Tools: Snakemake, Nextflow (fundamentals), BioPython, Pandas, NumPy, Jupyter, Flask
Data Engineering & Databases: SQL, ETL/ELT, API pipelines, Neo4j knowledge graphs, data modeling
Visualization: Plotly, R Shiny, Jupyter Dashboards, RMarkdown, interactive reporting

Experience

AI Safety Saarland — Interdisciplinary Research Incubator Nov 2025 – Present
AI Safety Research Fellow
- Developing **multi-omics ML models** for personalized oncology using fusion and graph-based approaches.
- Assessing **uncertainty, causality, and robustness** under cohort and distribution shifts.
- Building reproducible **Snakemake + Docker** pipelines with AI-safety methods for reliable benchmarking.

Myria Biosciences Jun 2024 – Present
Machine Learning & Bioinformatics Engineer
- Processed and analyzed **10,000+ genomic and proteomic samples**, reducing pipeline runtime by **35%** using workflow managers and HPC clusters.
- Benchmarked variant calling pipelines, improving SNP detection sensitivity by **8%** vs reference datasets.
- Built **ETL workflows (Python, Pandas, SQL)** to manage **50 GB/week** multi-omics data with **<2% error rate**, enabling efficient data onboarding.
- Engineered a **Neo4j knowledge graph** with **10,000+ biological entities**, improving R&D query speed by **50%**.
- Delivered **30+ interactive dashboards and reproducible reports**, accelerating research decisions across **3 clinical programs**.
- Developed and containerized reproducible NGS and ML pipelines using **Docker, Git, CI/CD**, increasing workflow reliability and automation.

Universität des Saarlandes Oct 2024 – Jul 2025
Research Assistant
- Applied **AI-driven CADD methodologies** for biosynthetic pathway optimization and molecule screening.
- Integrated synthetic biology with computational chemistry workflows for rapid in-silico experimentation.
- Automated ML pipelines using **Python, PyTorch, and Scikit-learn** on Linux, improving model testing and reproducibility.

Helmholtz Institute (HIPS) Jun 2024 – Dec 2024
Data Scientist in Drug Bioinformatics
- Analyzed and integrated NGS and **LC-MS/MS proteomics** datasets, contributing to the discovery of **15 novel biosynthetic gene clusters**.
- Developed ML models increasing expression efficiency prediction accuracy by **22%**.
- Applied **graph-based data science** for pathway discovery; **3 computational predictions validated experimentally**.
- Built Snakemake pipelines with version control, saving **40+ hours/month** of manual processing time.

- Delivered reproducible bioinformatics pipelines now used by multiple research groups for genomic and biochemical analysis.

KIST Europe Dec 2022 – Sep 2024
Research Assistant
- Supported sequencing-driven synthetic biology projects, validating **100+ molecular cloning experiments**.
- Led student training and workflow standardization, reducing onboarding time by **30%**.
- Executed and analyzed **ELISA, cloning workflows, and phage–host interaction** studies to support bacteriophage engineering.

Awards & Hackathons

Winner — Health Hack Saar 2025 *Saarbrücken, Germany*
MEDPlex: Bridging the Gender Data Gap in Pediatrics
- Built a real-time **evidence-based pediatric medication recommendation system** addressing gender bias in dosing.
- Delivered a full working prototype within 48 hours with potential to improve medication safety and reduce dosing errors.

Participant — Defensive Acceleration Hackathon 2025 *Global*
BioCast AI: Global Disease Outbreak Forecasting
- Developed a hybrid **LSTM + XGBoost** epidemiological forecasting pipeline with PCA/clustering analytics and 2025 outbreak risk predictions.

Publications

Reprogramming Filamentous FD Viruses to Capture Copper Ions. *ChemBioChem.* DOI: 10.1002/cbic.202400237 — 2024

Bacteriophage Engineering for Improved Copper Ion Binding. *Macromolecular Bioscience.* DOI: 10.1002/mabi.202300354 — 2023

Electrocardiogram Data Visualization and Dimensionality Reduction. *ICBHS 2022* — 2022

Brain Tumor Segmentation and Detection on MRI Images. *ICBHS 2022* — 2022

Projects & Certifications

- **GeneCoViz** — Co-expression network visualization tool (Plotly, NetworkX)
- **NeuroVision Explorer** — MRI analysis tool using CNNs and transformers
- **Heart Disease Prediction** — ML model achieving **92% accuracy** (Random Forest, XGBoost)
- **Neo4j Knowledge Graph** — Integrated 10,000+ biological entities linking proteomic and genomic metadata
- **FastAPI CI/CD Reference** — Production-ready FastAPI template with automated Docker-based CI/CD pipelines
- **Neo4j Certified Professional** — GraphAcademy
- **Neo4j Graph Data Science Certification** — GraphAcademy
- **ETL and ELT in Python** — (Short course)
- **AWS Cloud Practitioner Essentials** — AWS Training & Certification
- IBM, Stanford, Harvard: Deep Learning, AI with Python, Data Analysis, ML Pipelines

Education

Saarland University Oct 2022 – Present
M.Sc. Bioinformatics
- Thesis: "ML-Driven Exploration of NRPS Engineering in *E. coli*"
- Focus: Computational Biology, Genomics, NGS, Machine Learning

UET Lahore Oct 2018 – Aug 2022
B.Sc. Biomedical Engineering
- Graduated with Honors (GPA: 3.721/4.0)
- Thesis: "ECG Analysis - Monitoring of Signs & Symptoms"