Khushi Malik

Email: khushimalik511263@gmail.com | +91- 9536110472

LinkedIn | GitHub | Portfolio

EDUCATION

Bachelor of Science in Mathematics, **Kirori Mal College, University of Delhi**　　　*August 2023 – May 2027*
Current CGPA: *8.27/10.0*　　　*Delhi, India*

CORE COMPETENCIES

➢ **Programming & Data**: Python, Machine Learning, SQL Deep Learning, Natural Language Processing (NLP), Computer Vision, Data Analysis, Exploratory Data Analysis (EDA), Matplotlib
➢ **Tools & Frameworks:** Jupyter Notebook, Git, GitLab, PyTorch, AWS, Docker, FastAPI
➢ **Focus Areas:** Data Science
➢ **Specialized Skills:** Retrieval Augmented Generation (RAG), Model Fine-tuning, Prompt Engineering

PROFESSIONAL EXPERIENCE

Lead AI Researcher — Planto AI　　　*Remote, New Delhi ,　August 2025 – Present*
→ **Stremlined Retrieval-Augmented Generation (RAG) pipelines**, elevating Large Language Model accuracy by **18%** and reducing response latency by **25%** through advanced tokenization and prompt compression techniques.
→ **Led a cross-functional team of 5 engineers** in building an AI-powered VS Code extension, increasing developer productivity by **30%** through natural language coding assistance.
→ **Fine-tuned Large Language Models** using HuggingFace Transformers and PEFT methods, **cutting** computational costs by 40% while maintaining model performance.
→ **Presented technical findings** to stakeholders, contributing to strategic decisions that accelerated AI project adoption by **15%**.

AI Researcher — Planto　　　*Remote, New Delhi ,　May 2025 – July 2025*
→ **Researched and applied tokenization optimization and prompt compression** for Large Language Models, improving model inference speed by **20%** and increasing response relevance by **15%**.
→ **Engineered** end-to-end ML pipelines using Python and HuggingFace, enabling **production deployment within 4 weeks** and supporting **real-time data processing for 10K+ queries per day**.
→ **Collaborated** to **deploy** RAG systems, **enhancing** retrieval accuracy by 18%

Freelance Data Scientist — Zindi　　　*Remote, South Africa ,　April 2025 – Present*
→ Demonstrating expertise in **Computer Vision and Machine Learning** applied to large-scale agricultural datasets.
→ **Processed and engineered** features on 10,000+ samples, **boosting** model accuracy by 15% and **reducing** training time by 20%.
→ **Refined** ML models for agriculture and healthcare, **attaining** up to 92% predictive accuracy.

PROJECTS

Cogni-Compliance: End-to-End Agentic RAG System　| **GitHub** | **Live Demo** |**Documentation**　　September 2025
AI Chat Application for Legal & Regulatory Documents

→ Built an agentic RAG pipeline with FAISS + BM25 hybrid retrieval and cross-encoder re-ranking for accurate Q&A over GDPR, CCPA, and HIPAA.
→ Designed an agentic workflow with LangGraph, integrating Llama 3.1 for grading and Mixtral for answer synthesis.
→ Developed and deployed a full-stack MERN + FastAPI system, with RAG backend on AWS EC2 and frontend/backend on Render.
→ Processed legal PDFs (extraction, cleaning, semantic chunking, vectorization) using Colab/Jupyter to create a FAISS knowledge base.

Codex-Finetune: Intelligent Coding Copilot　| **Link**　　　*June 2025*
→ Fine-tuned starcoder 2-3b model on CodeSearchNet dataset using Low-Rank Adaptation (LoRA) techniques
→ Implemented multi-task code generation for functions, docstrings, and code completion using HuggingFace Transformers
→ Secured final training loss of 1.20 across 3,000 training steps with Weights and Biases monitoring
→ Tech Stack: Python, HuggingFace, Parameter Efficient Fine-Tuning, Machine Learning

Amini GeoFM Decoding the Field Challenge | Documentation
Crop Classification from Satellite Time-Series Data　　　*June 2025 – July 2025*

→ Engineered hybrid features by combining statistical and spectral vegetation indices with deep time-series embeddings from a PatchTST model.
→ Upgraded an ensemble classification pipeline integrating XGBoost, Logistic Regression, and SVM for robust multi-crop prediction (cocoa, rubber, oil palm).
→ Achieved Top 10% rank (10/100 teams) on Zindi leaderboard with a private log loss of 0.7473.
→ Generated final outputs as predicted crop type labels and probabilities, delivered in CSV format for analysis and submission.

AI Coder Pro: Visual Studio Code Extension |Link　　　*July 2025*
→ Built AI-powered Visual Studio Code extension using TypeScript and Node.js with Together AI integration
→ Implemented multi-agent tools for code review, bug detection, refactoring, and automated test generation
→ Created chat interface with memory capabilities and file upload functionality for enhanced user experience
→ Tools Used.: TypeScript, Node.js, Artificial Intelligence, Natural Language Processing.

CERTIFICATIONS & ACHIEVEMENTS

➢ Won prize money of USD 500 for most efficient solution.
➢ Earned Top 10% rank (10/100 teams)in the Amini GeoFM Decoding the Field Challenge
➢ Active contributor to open-source Machine Learning projects on GitHub
➢ Excel Business Intelligence – Link　　　SQL Beginner to Advanced – Link　　　Python Beginner to Advanced – Link