ABDULLAH KHAN

AI/ML Engineer | Machine Learning | Deep Learning | Generative AI

Phone: 03111823456 | **Email:** abdullahkhansherwani09@gmail.com

Location: Faisalabad, Punjab, Pakistan (Domicile: Punjab)

LinkedIn: linkedin.com/in/abdullahks-devhub | **GitHub:** github.com/abdullahks-devhub | **HuggingFace:** huggingface.co/spaces/AbdullahKS-Devhub

PROFESSIONAL SUMMARY

Final-year Computer Science student (CGPA: 3.57/4.0, Expected Graduation: May 2026) with hands-on experience across the full AI/ML lifecycle — from data preprocessing and feature engineering to model training, evaluation, and deployment. Built and deployed multiple production-grade ML applications using Python, Scikit-learn, and Streamlit on HuggingFace Spaces, integrating third-party APIs for real-time data retrieval. Completed specialized certifications from DeepLearning.AI and IBM in deep neural networks, hyperparameter tuning, Generative AI, agentic AI frameworks (LangGraph, CrewAI, AutoGen), and vector databases for RAG applications. Experienced in supervised and unsupervised learning, NLP, and prompt engineering, with exposure to LLMs and Generative AI application development. Passionate about designing intelligent, data-driven solutions and writing clean, well-documented, production-ready code.

EDUCATION

Bachelor of Science in Computer Science | CGPA: 3.57/4.0

The University of Faisalabad | *Expected Graduation: May 2026*

Punjab, Pakistan

TECHNICAL SKILLS

Programming Languages: Python, SQL, C, C++, Java, JavaScript, HTML5, CSS

Machine Learning & Deep Learning: Supervised & Unsupervised Learning, SVM, Artificial Neural Networks, CNNs, RNNs, Reinforcement Learning, Transfer Learning, Hyperparameter Tuning, Regularization & Optimization, Dimensionality Reduction, Anomaly Detection, Feature Engineering, Model Evaluation, Predictive Modeling, Applied Machine Learning

Generative AI & LLMs: Large Language Models (LLM), Prompt Engineering, Prompt Design, Retrieval-Augmented Generation (RAG), Vector Databases, Agentic AI Workflows, LangChain, LangGraph, LangGraph, CrewAI, AutoGen, BeeAI, Multimodal Prompting, Generative Adversarial Networks (GANs), AI Security, Responsible AI

NLP & Computer Vision: Natural Language Processing (NLP), Sequence Models, Convolutional Neural Networks (CNN), Image Analysis, Computer Vision

ML Libraries & Frameworks: Scikit-learn, TensorFlow, PyTorch, Keras, Pandas, NumPy, Matplotlib, Hugging Face

Data Science & Engineering: Data Preprocessing, Data Cleaning, Data Manipulation, Exploratory Data Analysis, Data Visualization, Data Structures, SQL, Relational Databases, Web Scraping, REST APIs, JSON, XML

Software Development: OOP, Software Deployment, Model Deployment, Flask, RESTful Web Services, Git/GitHub, Software Development Best Practices

Tools & Platforms: Streamlit, HuggingFace Spaces, Jupyter Notebook, PyCharm, MySQL, Oracle Database, XAMPP, Git, Pickle/Model Serialization

PROJECTS

CineMatch - Movie Recommendation System | *Python, Scikit-learn, Streamlit*

Live Demo | **GitHub Repository**

- Designed and implemented a content-based ML recommendation engine using cosine similarity on CountVectorizer-extracted feature vectors across a 5,000+ movie dataset, covering the full ML lifecycle from data preprocessing and feature engineering to model evaluation and deployment
- Applied data preprocessing and feature engineering techniques using Scikit-learn and Pandas to transform raw movie metadata (genres, cast, keywords, overviews) into optimized numerical representations for similarity computation
- Integrated the TMDB REST API to fetch real-time movie posters and metadata, demonstrating experience in connecting ML models with external APIs and backend data sources
- Serialized the trained model and similarity matrix using Pickle for efficient production loading, and deployed the application on HuggingFace Spaces for public access with cached model inference
- Wrote clean, modular, and well-documented Python code following software engineering best practices, with Git version control throughout development

HealthGuard AI - Multiple Disease Prediction System | *Python, Scikit-learn, Streamlit*

Live Demo | **GitHub Repository**

- Trained and deployed three independent SVM/Logistic Regression models for Diabetes, Heart Disease, and Parkinson's Disease prediction, managing the full ML lifecycle: data sourcing (UCI & Kaggle), preprocessing, feature scaling, model training, evaluation, serialization, and production deployment
- Engineered disease-specific preprocessing pipelines with dedicated StandardScalers for each model, applied across 1,266 patient records spanning 8–22 clinical features per disease, with models serialized as .sav files using Pickle for fast, dependency-free production loading
- Integrated trained ML models into a Streamlit web application with guided clinical input forms and colour-coded prediction outputs, demonstrating experience deploying models into user-facing applications and connecting ML backends to frontend interfaces
- Deployed to HuggingFace Spaces with @st.cache_resource for optimized model inference, modular code architecture for maintainability, and responsible AI disclaimers — reflecting production-grade software engineering and model monitoring awareness

CERTIFICATIONS

- **Machine Learning Specialization** - DeepLearning.AI & Stanford University
- **Deep Learning Specialization** - DeepLearning.AI
- **IBM RAG and Agentic AI Specialization** - IBM
- **Generative AI for Data Scientists Specialization** - IBM
- **Python for Everybody Specialization** - University of Michigan

ADDITIONAL INFORMATION

- Actively building an AI/ML portfolio with multiple production-deployed applications on HuggingFace Spaces, covering supervised learning, disease prediction, and recommendation systems
- Experienced in working with structured and unstructured data, applying NLP techniques, and integrating ML models with APIs and web applications for end-to-end intelligent systems
- Certified in Generative AI, agentic frameworks (LangGraph, CrewAI, AutoGen, BeeAI), vector databases for RAG, and deep neural network optimization through DeepLearning.AI and IBM
- Committed to writing clean, well-documented, and maintainable code with a strong interest in MLOps practices, model versioning, and deploying AI solutions that solve real-world business problems