

ABUZAR KHAN

+91-8329847784 abuzarkhan.ip123@gmail.com LinkedIn Portfolio GitHub

Summary

Data Scientist with expertise in Python, SQL, and Machine Learning, specializing in data-driven decision-making. Proficient in building predictive models, performing advanced analytics, and extracting insights from structured and unstructured data. Skilled in deep learning, NLP, and computer vision, with hands-on experience in deploying scalable AI solutions. Passionate about leveraging data science and AI to solve real-world challenges and drive business growth.

Education

MIT, Aurangabad

2021 – 2025

Bachelor of Technology (B.Tech) – Artificial Intelligence & Data Science

Maulana Azad College of Arts and Science

Higher Secondary Education (HSC)

Experience

Software Developer Intern

Apr 2025 – Oct 2025

NRB Bearings Limited

India

- Developed a GUI-based Tool Crib Management System using software engineering and database management skills.
- Built a web-based Kaizen Entry Dashboard leveraging frontend, backend, and data visualization skills.
- Applied AI and analytics skills to propose automation and optimization improvements in digital workflows.

Skills

Data Science / ML / DL: Regression, Classification, Clustering, Ensemble Methods, PCA, CNN, Model Tuning, Neural Networks, RNN/LightGBM, Feature Engineering

Gen AI: Groq LLM (Llama), Gemini 1.5, OpenAI GPT APIs, Hugging Face Transformers, n8n, RAG pipelines, Prompt Engineering

NLP: Text Preprocessing, Tokenization, Word2Vec, BERT, NLTK, spaCy, Text Classification

Data Analysis & Visualization: EDA, Matplotlib, Seaborn, Power BI, SQL (SQL Server), NoSQL (MongoDB)

Programming & Frameworks: Python (NumPy, Pandas, SciPy, Scikit-Learn, TensorFlow, Keras, PyTorch, FastAPI), React

MLOps & Deployment: CI/CD, Docker, Render, Git/GitHub, IIS services, Flask, Gunicorn

Projects

Credit Card Fraud Detection System | Python, Flask, scikit-learn, Random Forest

Link

- Built a Flask-based web application with REST API for real-time credit card fraud detection using Random Forest classifier.
- Developed user-friendly web interface and API endpoint for programmatic access to fraud predictions.
- Implemented model training pipeline with feature engineering and achieved reliable fraud detection accuracy.
- Deployed application ready for production on cloud platforms like Render with CI/CD integration.

E-Commerce Business Intelligence Dashboard | Power BI, SQL, DAX, Excel

Link

- Developed an interactive Power BI dashboard for e-commerce analytics providing key insights into sales, profit, customer trends, and payment modes.
- Utilized by business stakeholders for decision-making and sales forecasting.

Real-Time Facial Expression Detection Model | Python, OpenCV, TensorFlow/Keras

- Developed a CNN-based model for classifying facial expressions in real-time from video streams using FER-2013 dataset.
- Implemented real-time webcam data augmentation for improved model performance.
- Applied in emotion-aware systems for HCI, personalized AI interfaces, and mental health insights.

Named Entity Recognition (NER) for Product Data | Python, spaCy, PyTorch, OpenCV

- Automated entity recognition from product images to extract specifications including dimensions, weight, and voltage details.
- Achieved 78% precision in entity extraction, enhancing inventory management and e-commerce product listings.

Certifications

- Career Essentials in Data Analysis by Microsoft and LinkedIn
- Internship Common Aptitude Test by LABMATIC
- Machine Learning, NLP, MLOps & Deployment Bootcamp – Udemy (By Krish Naik)