MOHAMMAD SARFRAZ KHAN
Data Science & Data Analyst

📍 Sitamarhi, Bihar | 📞 +91 7463849811 | ✉ **sarfrazkhan026010@gmail.com** | 🔗 Linked In | 🔗 Github

SUMMARY

Detail-oriented Data Science & Machine Learning enthusiast with hands-on experience in Python, SQL, Power BI, Machine Learning, Deep Learning, NLP, and Computer Vision. Skilled in full ML lifecycle including data collection, preprocessing, model development, performance tuning, and deployment. Strong analytical mindset with the ability to translate data into actionable insights. Seeking an entry-level Data Analyst / Machine Learning Engineer role.

TECHNICAL SKILLS

Programming: Python, SQL
ML/DL: Scikit-Learn, TensorFlow, Keras, OpenCV, CNN, ResNet, NLP
Data Analysis: Pandas, NumPy, EDA, Statistics, Feature Engineering
Visualization: Power BI, Matplotlib, Seaborn
Tools: Git, Jupyter Notebook, Power Query, DAX
Database: MySQL

PROJECTS

OralScan AI – Oral Cancer Detection System (CNN + Transfer Learning)
Python | TensorFlow | Keras | OpenCV | Computer Vision | Image Classification
- Built a CNN + transfer learning model achieving 94% accuracy for oral cancer detection.
- Automated image preprocessing and augmentation to improve model robustness.
- Developed a real-time interface for early oral lesion detection to assist healthcare professionals.

Digit Recognition Using CNN (MNIST)
TensorFlow | Deep Learning | OpenCV
- Developed a CNN model with **98%+ accuracy** for handwritten digit recognition.
- Added dropout, batch normalization, and Adam optimizer to improve performance.
- Integrated OpenCV for real-time prediction from user-drawn input.

Patients Analytics Dashboard
Power BI | SQL | DAX
- Created an interactive dashboard for hospital operations and billing performance.
- Optimized DAX queries and Power Query transformations reducing loading time by **30%**.
- Implemented **Row-Level Security (RLS)** for data privacy compliance.

Sentiment Analysis Web App (Text & Voice)
Python | Streamlit | Scikit-learn | NLP | Machine Learning
- Built a sentiment analysis app classifying text and voice input as Positive, Negative, or Neutral.
- Implemented speech-to-text conversion and TF-IDF-based ML model for robust predictions.
- Developed a user-friendly interface for real-time sentiment detection and visualization.

EXPERIENCE

Data Science Intern – TechSimPlus Learning, Bhopal
June 2024 – August 2024
- Performed EDA, data preprocessing, and exploratory insights using Python & SQL.
- Applied Machine Learning, Deep Learning, NLP, and Computer Vision techniques on real datasets.
- Designed dashboards in Power BI to derive actionable business insights.

EDUCATION

B.Tech – Artificial Intelligence & Data Science (AIDS)
LNCT Group of Colleges, Bhopal
CGPA: 7.0 | 2022 – 2026

Higher Secondary (Class 12)
Bihar School Examination Board
61.5%

Sacred Heart School ,Sitamarhi(Class 10)
Central Board Secondary Education
58%