Ibrahim Zia Khan

ziaibrahim90@gmail.com
my-portfolio-ibrahim.vercel.app
github.com/Ibrahim-ziaa
+92-319-4447204
Upwork Profile

Education

- **National University of Computer and Emerging Sciences (FAST)** Lahore, Pakistan
  BS in Data Science *Aug. 2022 – May. 2026*

Experience

- **Freelance Data Scientist** Remote
  Upwork.com *Jan. 2024 – Present*
  - **Clientele**: Worked with clients from the USA, Singapore, Japan, Ireland, and Pakistan across domains like healthcare, marketing, and e-commerce.
  - **Projects**: Delivered 7+ successful jobs ranging from end-to-end ML pipelines to real-time dashboards and NLP tools.
  - **Technologies**: Utilized Azure Databricks, Power BI, HuggingFace Transformers, Streamlit, and PyTorch under Upwork Enterprise contracts.

Projects

- **AI-Based Skin Cancer Diagnostic System** Designed an intelligent diagnostic platform to detect skin cancer from lesion images using both deep learning and traditional ML approaches. The pipeline included preprocessing with the Dull Razor method for hair removal, k-means clustering for segmentation, and feature extraction using ABCD and GLCM metrics. EfficientNet-B0 handled image classification, while SVM and LightGBM were trained on handcrafted features. Explainability was achieved through Grad-CAM for CNN visualization and SHAP/LIME for tabular insights.
- **Twitter Sentiment Analysis and Visualization** Built a sentiment classification system using BERT and LSTM models on a corpus of tweets. Preprocessing included tokenization, noise filtering, and sequence padding. After classification, results were visualized using Power BI, where keyword clouds, sentiment trends, and location-based polarity breakdowns were presented. This tool was optimized for real-time dashboards for social media teams.
- **Osteoarthritis Detection via X-rays** Developed a model to detect signs of osteoarthritis in knee X-rays using ResNet50 for image classification. To improve generalization, data augmentation techniques like flipping, rotation, and brightness adjustment were used. The trained model was further enhanced by Grad-CAM visualizations, which allowed medical experts to validate the prediction focus areas.
- **Voice-to-Text Summarizer Platform** Built a full-stack Streamlit application that takes voice input and returns a clean, concise summary. For transcription, HuBERT and Wav2Vec models were employed, while summarization was handled using BART and Mistral API. The system was designed with both GPU and CPU runtime fallback and supports various audio formats including MP3, WAV, and FLAC.
- **Google Ads Performance Dashboard** Developed a Looker Studio dashboard integrated with Google Sheets and Google Ads API. The dashboard provided real-time insights on CTR, CPC, ROAS, and conversion rates. It featured automated anomaly detection, funnel visualizations, and daily email reports for marketing executives.

Certificates

- IBM Python Data ScientistCoursera Verified Certificate
- IBM Data VisualizationCoursera Verified Certificate

Skills

- LanguagesPython, R, SQL, C++
- TechnologiesPyTorch, TensorFlow, HuggingFace, Scikit-learn, OpenCV, Azure Databricks, Power BI, Tableau, Looker Studio, Streamlit, PostgreSQL, Firebase