Yousha Saibi

github.com/youshasm
youshasaibi187@gmail.com
linkedin.com/in/yousha-saibi
Islamabad, Pakistan
+92-3152424468

PROJECTS

- **Smart Traffic Management System**: YOLO-based system for smart signal control and road surveillance. **Technologies:** Python, OpenCV, YOLOv8/11, EasyOCR

  - **Training**: Annotated datasets via Roboflow to train YOLOv8 and YOLOv11 models.

  - **Detection**: Applied models on live video to detect vehicles, pedestrians, and traffic signals.

  - **OCR**: Extracted license plate numbers using bounding box cropping and EasyOCR.

- **Real-Time SemiStream Mesh Join Architecture**: An ETL architecture for ingesting and transforming live-stream data into a data warehouse. **Technologies:** Java, SQL.

  - **Semi-Stream Join**: Applied Mesh Join architecture to join stream data with master data in real-time using a hash table.

- **Reviews Sentiment Classifier**: An NLP-based system to predict sentiment polarity from customer reviews in online marketplaces. **Technologies:** Python, NLTK, Scikit-learn, Pandas.

  - **Preprocessing**: Cleaned and tokenized textual data, removed stop words, and applied stemming.

  - **Model Training**: Used Naive Bayes and Logistic Regression for sentiment classification.

- **Tuberculosis Data Analysis Dashboard**: An interactive dashboard visualizing TB trends across countries using d3.js. **Technologies:** JavaScript, d3.js.

  - **Visualizations**: Created sunburst charts, bar plots, and global maps for comparative insights.

EDUCATION

- **National University of Computer and Emerging Sciences** — Islamabad, Pakistan
  Bachelor of Science in Data Science — *Aug. 2022 – May. 2026*
  - **Relevant Courses**: Data Structures, Algorithms, Adv. Statistics, Artificial Intelligence, Operating Systems, Data Warehousing & Big Data Analytics.

PROGRAMMING SKILLS

- **Languages**: Python, SQL, R, C, C++, JavaScript.

- **Technologies**: PyTorch, Tensorflow, LangChain, Keras, Transformers, HuggingFace, MLLib, D3.js, Tableau, Plotly, Scikit-learn, Matplotlib, Pandas, NumPy, PySpark, Hadoop, Flask, MongoDB, Docker and PostgreSQL