KABIR SHAH

kabir.js187@gmail.com | https://www.linkedin.com/in/kabirjshah/ | 857-384-8674

SKILLS

- ***Languages***: Java, Python, C, C++, HTML/CSS, JavaScript
- ***Frameworks and libraries***: Django, Flask, FastAPI, Pandas, NumPy, Scikit-learn, ADABoost, XGBoost, TensorFlow, Keras, PyTorch, spaCy, PySpark, Matplotlib, Plotly, Seaborn, Hadoop, PyPDF2, pdfplumber, OpenCV, dlib, pytesseract
- ***Databases***: MySQL, MongoDB
- ***Cloud-platforms***: AWS, GCP, Azure
- ***Analytical and ML:*** Regression, Classification, Clustering, Cross-Validation, Large Language Models, Computer Vision

EXPERIENCE

Jr Data Scientist | *Brainer Hub Solutions, Ahmedabad, India* **Jun'23 – Dec'24**

- Developed and maintained advanced data processing systems for diverse projects—including an electric & gas bill automation tool, an NLP-based medical report extraction system for a biotech firm, and a university analytics portal—resulting in a 40% reduction in data errors and a 35% boost in processing throughput.5%.
- Designed and deployed an NLP-driven document parsing framework for high-volume medical reports (processing 10,000+ documents/month), leveraging named entity recognition, classification, and pattern-based extraction to enhance clinical tagging accuracy by 30% and reduce manual review time by 60%.
- Implemented robust automated testing frameworks and CI/CD pipelines to validate model performance and data integrity, collaborating within agile teams to expedite deployment cycles by 20% and elevate system reliability by 30%.

Data Science Intern | *Brainer Hub Solutions, Ahmedabad, India* **May'23 - Jun'23**

- Engineered scalable data ingestion and analytics pipelines for a real estate and social engagement platform using Django, Flask, FastAPI, and pandas; optimized operational analytics by 60% and improved platform responsiveness by 25%.%.
- Integrated a high-precision facial recognition analytics module for corporate security enhancements using TensorFlow, Keras, and PyTorch to extract actionable features, boosting recognition accuracy by 30% and reducing processing latency by 35%.
- Coordinated iterative feature enhancements with rigorous documentation and testing protocols, refining model outputs and data processing pipelines to achieve a 20% reduction in deployment time and enhanced analytical precision.

Data Science Intern | *E-Seller Hub, Ahmedabad, India* **Jan'23 - Apr'23**

- Developed robust Python-based data pipelines for real-time ingestion, transformation, and predictive modeling across multi-channel retail operations, increasing throughput by 20% and shortening decision latency by 15%.
- Implemented cloud-native deployments on AWS, Azure, and GCP, integrating advanced statistical models with automated error-handling frameworks that reduced manual intervention by 45% while ensuring scalable data processing.

PROJECTS

Unique ID Management System | *Technologies: Python, MySQL, HTML, CSS, JavaScript*

- Engineered a robust identity management platform featuring multifactor authentication and comprehensive data logging to enable advanced user behavior analytics, boosting verification accuracy by 30% and reducing processing time by 25% for over 1,000 users.

EDUCATION

Northeastern University, *Boston, MA* **Jan'25 - May'27**
Master of Science in Computer Science

Gujarat Technological University, *Gandhinagar, India* **Aug'19 - May'23**
Bachelor of Engineering - Computer Engineering