Shraiy Khaddar

Machine Learning Engineer

✉ shraiykhaddar23@gmail.com

📞 +91 9171656905

📍 Hyderabad, Telangana, India

🔗 linkedin.com/in/shraiykhaddar

🔗 github.com/Shraiy

🔗 https://shraiykhaddar.netlify.app/

Education

09/2021 – present Gwalior, India	**Bachelor of Technology** *Madhav Institute of Technology and Science* GPA: 7.52
2020 – 2021 India	**Senior Secondary** *Navankur Higher Seconday (MPBSE)* Percent: 76.8
2018 – 2019 India	**Secondary** *Samarpan Public Senior Secondary (CBSE)* Percent: 84.5

Professional Experience

12/2024 – 06/2025 Gwalior, India	**Data Science Intern** *Saanvi Innovative Solutions* • Researched 50+ AI/ML papers and developed 15+ models (CNNs, RNNs, LLMs) with up to **99% accuracy** across various tasks. • Preprocessed **20M+ records** and **10K+ images**, applying **TF-IDF**, lemmatization, and **SMOTE** to enhance data quality and model balance. • Built a **multiclass sentiment analysis system** on **2M+ product reviews** using Random Forest, Naive Bayes, **DistilBERT**, and **RoBERTa**, driving insights for **demand forecasting and product strategy**. • Visualized sentiment trends and model metrics using **Matplotlib**, **Seaborn**, and WordClouds; collaborated on integration into decision-making pipelines.

Projects

05/2024	**Fraud Transaction Detection** 🔗 • Developed an AI-based fraud detection system on 7M+ transactions with 11 attributes. • Technical Skills Conducted EDA and applied SMOTE for class balancing. • Trained models (XGBoost, Random Forest, Logistic Regression), achieving 99% accuracy. • Implemented data visualization techniques to analyze fraud patterns. • Deployed the user ready interface on Render, enabling real-time fraud detection.
02/2024 – 04/2024	**Deepfake Detection Using CNNs** 🔗 • Fine-tuned a deepfake detection model using MesoNet and other CNN models with 3+ layers. • Preprocessed 12,000+ images, applying image augmentation & normalization. • Optimized and trained CNN model using binary cross-entropy, achieving 90%+ accuracy. • Evaluated results using confusion matrix, classification report, and F1-score.

09/2023 – 12/2023	**Accident Severity Prediction** ⬈

- Built an ML model using 7M+ records & 46 features to predict accident severity.
- Performed data preprocessing, feature extraction, and **NLP**-based text processing (TF-IDF vectorization) on 3M+ instances.
- Applied **SMOTE** to address the class imbalance increasing 2M+ instances.
- Implemented Random Forest, XGBoost, and other ML models to achieve 90%+ accuracy, by evaluating precision-recall, and feature importance analysis.
- **Tech Stack:** Python (Pandas, NumPy, Scikit-learn, Imbalanced-learn, NLTK), ML Algorithms, Feature Engineering.

Technical Skills

Languages
Python, C++, SQL, MySQL, HTML
Data Structures and Algorithms

Machine Learning
XGBoost, Neural Networks (CNN), LSTMs, NLP (LLM)

Data Science
- Feature Engineering, EDA, Data Cleaning, TF-IDF

Tools and libraries
Libraries: Pandas, Numpy, Matplotlib, Tenorflow
Tools: Git, Jupyter, Google Colab, VS Code

Certificates

Data Analytics with Python (NPTEL) ⬈
Score: 73%

Big Data Computing (NPTEL) ⬈
Score: 75%

Responsible and Safe AI Systems (NPTEL) ⬈
Score: 73%

Accomplishments

- **Completed Summer Internship in MySQL & RDBMS** (*Secured 'A' Grade*).
- **Certificate of Participation** in 39th M.P. *Young Scientist Congress Innovation Challenge.*
- **Active participant** in college tech events & AI/ML workshops.