

# Shraiy Khaddar

## Machine Learning Engineer



✉ shraiykhaddar23@gmail.com

☎ +91 9171656905

📍 Hyderabad, Telangana, India

🌐 [linkedin.com/in/shraiykhaddar](https://www.linkedin.com/in/shraiykhaddar)

🐙 [github.com/Shraiy](https://github.com/Shraiy)

🔗 <https://shraiykhaddar.netlify.app/>

### Education

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

### Professional Experience

12/2024 – 06/2025 Gwalior, India	<b>Data Science Intern</b> <i>Saanvi Innovative Solutions</i> <ul style="list-style-type: none"><li>• Researched 50+ AI/ML papers and developed 15+ models (CNNs, RNNs, LLMs) with up to <b>99% accuracy</b> across various tasks.</li><li>• Preprocessed <b>20M+ records</b> and <b>10K+ images</b>, applying TF-IDF, lemmatization, and SMOTE to enhance data quality and model balance.</li><li>• Built a <b>multiclass sentiment analysis system</b> on <b>2M+ product reviews</b> using Random Forest, Naive Bayes, <b>DistilBERT</b>, and <b>RoBERTa</b>, driving insights for <b>demand forecasting and product strategy</b>.</li><li>• Visualized sentiment trends and model metrics using <b>Matplotlib</b>, <b>Seaborn</b>, and WordClouds; collaborated on integration into decision-making pipelines.</li></ul>
-------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Projects

05/2024	<b>Fraud Transaction Detection</b> <a href="#"></a> <ul style="list-style-type: none"><li>• Developed an AI-based fraud detection system on 7M+ transactions with 11 attributes.</li><li>• Technical Skills Conducted EDA and applied SMOTE for class balancing.</li><li>• Trained models (XGBoost, Random Forest, Logistic Regression), achieving 99% accuracy.</li><li>• Implemented data visualization techniques to analyze fraud patterns.</li><li>• Deployed the user ready interface on Render, enabling real-time fraud detection.</li></ul>
02/2024 – 04/2024	<b>Deepfake Detection Using CNNs</b> <a href="#"></a> <ul style="list-style-type: none"><li>• Fine-tuned a deepfake detection model using MesoNet and other CNN models with 3+ layers.</li><li>• Preprocessed 12,000+ images, applying image augmentation &amp; normalization.</li><li>• Optimized and trained CNN model using binary cross-entropy, achieving 90%+ accuracy.</li><li>• Evaluated results using confusion matrix, classification report, and F1-score.</li></ul>

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.