

Masihullah Shaik

+91 7674026786 | masihulla17@gmail.com | LinkedIn | GitHub | Portfolio

EDUCATION

Indian Institute of Information Technology Sri City

India

B.Tech (Honors) in Computer Science and Engineering, CGPA - 9.41/10

Aug 2018 - May 2022

EXPERIENCE

Swiggy, Data Scientist 2

Jun 2022 - Present

- Joined as *Data Scientist 1* in Jun 2022 and promoted to *Data Scientist 2* in Oct 2023.
- Refund Fraud Detection:** Developed an **E2E fraud claims detection pipeline** using ensemble of TabNet(classification) and DevNet(anomaly detection). Handled limited labeled data (5%) and class imbalance (10% positives) with weak and semi-supervised methods. Engineered over 250 features to **detect 70+ fraud patterns** with 85% precision.
- Customer Lifetime Value:** Formulated a **predictive unified customer score** basis 200+ user attributes by estimating it as function of lifetime value and potential risk. Trained a multi-output neural network to rank 5M new users with minimal order data, **addressing information sparsity and cold start issues**.
- Menu Enhancement:** Given a dish name, predict whether its veg or non-veg, and recommend a dish category. Built **multi-head attention neural network** using N-grams and Word2Vec embeddings, achieving 96% precision & reducing manual correction tickets by 75% (~10K per day).
- Text-to-SQL using Gen AI:** Contributed to an in-house **RAG-based Text-to-SQL workflow** enabling users to ask questions in natural language and receive SQL queries, used by 1,000+ employees. Performed related-work industry analysis, anecdotal analysis and proposed new layers in the workflow to improve the accuracy.
- Restaurant Location Correction:** Designed and implemented a framework to rectify incorrectly logged restaurant locations using delivery partners' **historical GPS data** and the **Geometric Median** technique. Achieved 92% accuracy, correcting over 13K locations and improving delivery efficiency.

Data Science Intern

Jun 2021 - Jan 2022

- Fraud Rings Detection:** Designed a **graph-based anomaly detection framework** using the Leiden community detection algorithm, enabling real-time identification of fraud groups. Improved F1 score by 9.92% compared to existing SOTA like GCN and DMON.

IBM, GRM Research Intern

Jan 2020 - Jun 2020

- Road and Pothole Segmentation:** Integrated **Attention-based Refinement and Feature Fusion** modules into the DeepLabV3+ architecture to enhance spatial representation and global context. Achieved a mIoU of 93.6% for road and 73.83% for pothole segmentation, with an 81% reduction in run-time.

Vsualthree60, Machine Learning Intern

Jul 2019 - Jan 2021

- AI SaaS Platforms:** Developed platforms for Crowd Surveillance, Image Docs Parsing, Appointments Chatbot, and Virtual Clothing Try-On. Used deep learning models like **VGG-Face, YOLO, Tesseract, OpenCV, and Rasa** for age, gender, and ethnicity prediction, OCR in Arabic and English, and image/text processing.

PUBLICATIONS

- Utilizing DevNet with Variational Loss for Fraud Detection in Hyperlocal Food Delivery.** Published in CODS-COMAD 2024, Bangalore, India. [Paper Link].
- Identifying Fraud Rings Using Domain Aware Weighted Community Detection.** Published in CD-MAKE 2022, Vienna, Austria. [Paper Link].
- A Decentralized Collaborative Strategy for PTZ Camera Network Tracking System using Graph Learning.** Published in AMMS 2022, Paris, France. [Paper Link].
- Attention Based Coupled Framework for Road and Pothole Segmentation.** Published in ICPR 2020, Milan, Italy. [Paper Link].

TECHNICAL SKILLS

- Programming & Tools:** Python, SQL, PySpark, Pandas, Numpy, Tensorflow, Keras, Pytorch, OpenCV, Sklearn
- Others:** LLM, RAG, Stable Diffusion, Prompting, Databricks, Snowflake, Kafka, AWS, GCP, Git, Django, Flask

RECOGNITION

- Most Valued Player:** Awarded 4 times for **problem solving abilities** and **remarkable impacts**. @Swiggy
- Swiggstar Award:** For demonstrating **Never Settle** attitude. @Swiggy