Hardik Soni

C-802, Metrojazz, Mahalunge, Pune, Maharashtra 411045

+91-8805020774 | iamhardikat11@gmail.com | linkedin/hardik-soni | github.com/hs094 | hs094-Portfolio

Education

Indian Institute of Technology, Kharagpur **Dec 2020 – May 2025**
Integrated Bachelors and Masters of Technology in Computer Science *7.74 / 10*

M.P Junior College **Apr 2018 – Apr 2020**
Maharashtra State Board of Secondary and Higher Secondary Education *85.58%*

Awards and Achievements

- Secured an **AIR 400** in **JEE Advanced** among 250,000 candidates and **AIR 1952** in **JEE Mains** among 1.2 million candidates.
- Received **Reliance Foundation Scholarship in Artificial Intelligence and Machine Learning** among 5000+ candidates.
- Achieved a maximum rating of **1551 (Specialist)** on **Codeforces** and **1973 (Guardian)** on **LeetCode** under handle **hs094**.

Experience

Axtria Inc. | *Software Developer Intern* | *Smart Data Ingestion System* **May 2024 – July 2024**
Objective: Designed a scalable data ingestion system with automated format detection and cloud-native streaming

- Developed a robust rule-based parser for **format detection**, enhancing data pre-processing accuracy and reducing ingestion faults.
- Implemented a memory-optimized **preview engine** for large datasets, supporting dynamic delimiters, column and header settings.
- Built a unified **cloud ingestion layer** for **AWS S3 Buckets**, **Azure Blob**, enabling ZIP file parsing and interactive previews.
- Reduced ingestion failures by **30%** by proactively debugging and optimizing the **DataMAx Catalog** import/export pipeline.

Brigham Young University | *Research Intern – Human-Centered Security* | *UPI Fraud Detection* **Oct 2022 – July 2023**
Objective: Analyzed UPI fraud behaviors and recommended secure design strategies based on user research.

- **Built** a fully functional **UPI simulator** in **Swift (iOS)** to emulate real-world payment fraud scenarios for realistic threat modeling.
- Conducted a **mixed-methods study** with **24 participants** across age groups to assess behavioral risks, revealing friction points in trust and error handling.
- **Applied** statistical tests, including **Mann-Whitney U**, **Kruskal-Wallis**, and **logistic regression**, to identify fraud predictors.
- **Synthesized findings** into **design suggestions** for UPI apps to enhance **fraud resilience**, **user trust**, and **transaction clarity**.

Projects

Credit Fraud Detection System | *XGBoost, LightGBM, Imbalanced Learning* **Feb 2024 – March 2024**
Objective: Developed a **high-precision fraud detection system** processing **284,807 transactions** with **98.4% AUC**

- Implemented **advanced techniques for imbalanced datasets**, addressing rare fraud events ($< 0.2\%$ of transactions) through strategic resampling and algorithmic adjustments, improving fraud capture rate by **37%**.
- Engineered a **robust cross-validation framework** with 5-fold validation, achieving consistent performance across folds with scores ranging from **94.8% to 99.4% AUC**.
- Optimized model training pipeline with early stopping mechanisms that reduced training time by **40%** while maintaining accuracy.

H&M Personalized Fashion Recommender | *TensorFlow, Two-Tower Vector Networks, MLOps* **Oct 2023 – Dec 2023**
Objective: Developed a **real-time recommendation system** processing **31M+ transactions** across **105K fashion articles**

- Engineered **feature pipelines** using Polars and Sentence Transformers, creating **semantic embeddings** and **temporal features** capturing seasonal fashion trends.
- Architected a **4-stage recommendation system** (candidate generation, retrieval, ranking, re-ranking) delivering **personalized recommendations** with **sub-second latency**.
- Deployed **MLOps infrastructure** with Hopsworks feature store and Kubernetes-based inference services, reducing **model deployment time** by 60%.

Relevant Coursework

Theory + Lab: Operating Systems, Computer Networks, Database Management Systems, Computer Organisation & Architecture, Compilers, Software Engineering, Programming & Data Structures, Algorithms-I & II, Distributed Systems
Theory: Deep Learning, Machine Learning, Probability & Statistics, Statistical Inference, Discrete Structures, Linear Algebra

Technical Skills

Languages: Python, C/C++, SQL, Bash, LaTeX, MIPS, Assembly
Skills: Machine Learning, Recommender Systems, Feature Engineering, MLOps, Data Science, Systems Programming
Technologies/Frameworks: TensorFlow, Keras, CatBoost, XGBoost, LightGBM, Polars, Pandas, NumPy, scikit-learn, Sentence Transformers, Hopsworks, Kubernetes, Flask, Git

Extracurricular Activities

Student Mentor | *Students' Welfare Group, IIT Kharagpur* **Dec 2022 – Present**