

YAMIN HOSSAIN

Yamin's Portfolio

019-779-40357

issan.yamin@gmail.com

Kaggle Profile

Yamin Hossain

YaminH

Professional Summary

Data Scientist with a foundation in Computer Science and a specialization in Data Analytics. Demonstrated expertise in building and deploying machine learning models, with hands-on experience in Natural Language Processing, computer vision, and RAG-powered systems. Eager to apply skills in deep learning and model optimization to solve complex business problems.

Technical Skills

Programming Languages: Python, SQL, R, Java, MATLAB, HTML/CSS, JavaScript, LaTeX

Data Science & ML: TensorFlow, PyTorch, Scikit-Learn, NLP, Transfer Learning, AutoGluon

Databases: MySQL, Microsoft SQL Server, PostgreSQL, MongoDB, Pinecone, Cassandra, Redis, Neo4j

Tools and Frameworks: Flask, Langchain, Apache Hadoop, Streamlit, Tableau, Power BI, Looker Studio, Docker, Google Cloud Platform.

Recent Projects

Art Intelligence Hub: Ensemble Deep Learning for Art Analysis | *TensorFlow, Keras, Grad-CAM* **August 2025**

- Launched **Art Intelligence Hub**, a publicly accessible, multi-page Streamlit application featuring dual deep learning classifiers for art authenticity and style identification.
- Constructed a high-accuracy **ensemble model** for art style classification by averaging predictions from fine-tuned **EfficientNetV2** and **ConvNeXt** models on a dataset of over 150,000 images.
- Integrated model interpretability features, including **Grad-CAM** visualizations and detailed prediction breakdowns, to enhance user trust and model explainability.

Cortex AI: Multi-Model Insights Hub | *LangChain, ChatGroq API, LLM, Streamlit* **January 2025**

- Developed a RAG-powered platform integrating **multiple LLMs (DeepSeek, Qwen, Gemma, LLaMA, Mistral)** for document analysis, **achieving 98% retrieval accuracy and reducing response latency by 40%**.
- Engineered a high-performance vector retrieval system using **LangChain's DocArrayInMemorySearch**, enabling real-time context-aware responses, multi-model comparisons, and Dynamic data visualizations generated from LLM outputs.

ReelSenseAI | *Flask, TMDb API, Render, LLaMA3* **April 2024**

- Created a **Flask-based** movie and TV show recommendation app leveraging the **TMDb API**, enabling users to browse by genre and receive recommendations generated via an LLaMA 3-integrated chatbot.
- Integrated **LLaMA 3 and Groq API** for interactive movie recommendations, reducing query **response time by 35%** and enhancing accuracy in personalized recommendations.

Experience

Code Casa **Feb 2024 – Mar 2024**

Data Science Intern

Remote

- Implemented ML models for **email spam and fraud detection**, achieving **97% precision** and **reducing false positives by 22%** using ensemble classifiers and feature selection techniques.
- Designed CNN-based gender and age prediction models, **achieving 92% accuracy**, **reducing classification errors by 12%**, and **optimizing inference time by 20%**.

Prodigy InfoTech **Jan 2024 – Feb 2024**

Machine Learning Intern

Remote

- Engineered a house price prediction model (**R²: 0.87 on test data**) using regression techniques.
- Implemented deep learning systems for **hand gesture (accuracy: 90%)** and **food item classification (accuracy: 88%)**, enhancing interaction and tracking accuracy.

Education

Vellore Institute of Information Technology, Andhra Pradesh **2021 – 2025**

Bachelor of Technology in Computer Science and Engineering (Data Analytics)

CGPA: 8.14

Relevant Coursework: Data Warehousing & Data Mining, Mining Massive Datasets, NoSQL Databases, Optimization Techniques, Data Structures & Algorithms, Database Management Systems.

Noubahini College, Dhaka **2020**

Higher Secondary School Certificate, Intermediate and Secondary Education Boards, Bangladesh

CGPA: 5(Scale of 5)