

Kaustubh Rath

Email: 1.0manydreams@gmail.com | Phone: +91-7499718082
GitHub: github.com/Kaustubh-Rathi | Codeforces: Pathu123 | LinkedIn: kaustubh-rathi-9228ab255 | Personal Website: kaustubh-rathi.github.io

Education

B.Tech in Information Technology (Minor in Economics), Indian Institute of Information Technology (IIIT), Allahabad July 2022 - June 2026
CGPA: 8.52/10 (as of 4th Year)
Secured **99.45 percentile** in JEE Mains 2022 (Top 0.65% nationally)

Experience

ASPR AGROVALUE (I) LIMITED — AI Intern Dec 2024 - Feb 2025

- Enhanced Qwen-1.5B model through LoRA fine-tuning on Google Cloud Vertex AI, reaching **85% accuracy** in farmer procurement predictions, leading to a **20% reduction in wasted resources**.
- Built agentic AI system for real-time data scraping/processing (**1,000+ records/day**), improving operational efficiency by **30%**.
- Integrated RAG with ensemble models (XGBoost/Random Forest) using 5+ data sources to enhance prediction reliability.
- Engineered a Named Entity Recognition (NER) system using Python and NLP techniques (RAG, SLMs); system automated extraction of key data points, processing **10,000+ documents monthly with 95% accuracy**.

Projects

AI-Powered Multilingual Clinical Interview Chatbot [GitHub link - [🔗](#)]
Challenge: Limited access to clinical interviews due to language barriers.
Outcome:

- Architected multilingual NLP pipelines with Langchain and FAISS, spanning 5+ languages, to amplify clinical reach.
- Fused symptom extraction and automated reporting via TensorFlow, securing 89% precision.

Movie Magic [GitHub link - [🔗](#)]
Challenge: Personalized movie recommendations based on user preferences.
Outcome:

- Developed a KNN model on 5,000+ IMDB records using scikit-learn, generating 2,000+ monthly recommendations.
- Integrated search bar with advanced filters, increasing user browsing time by 20% and boosting satisfaction.

Federated Learning for Image Classification on Raspberry Pi [GitHub link - [🔗](#)]
Challenge: Distributed model training on low-power devices.
Outcome:

- Trained MobileNetV2 via federated learning on 2+ Raspberry Pi clients using 500+ images.
- Achieved 92% accuracy and 80% server load reduction; deployed model on Edge Impulse.

End-to-End Wine Quality Prediction Pipeline [GitHub link - [🔗](#)]
Challenge: Subjective quality evaluation through manual tasting.
Outcome:

- Crafted an ML pipeline with TensorFlow, yielding a **0.45 mean absolute error** and **0.85 R-squared** across **1,600+ samples**.
- Streamlined access for **50+ users** by deploying a Flask application on AWS, accelerating usability by **30%**.

Technical Skills

- Programming Languages:** Python, C++
- AI & Machine Learning:** TensorFlow, PyTorch, Scikit-learn, Langchain, Regression, Classification, PCA, Random Forest, XGBoost
- Data Analysis & Visualization:** NumPy, Pandas
- Tools & Technologies:** Google Cloud Platform (Vertex AI), Docker, Git, Unix/Linux, TCP/IP, MySQL, MongoDB, Apache Spark

Related Coursework

Data Structures and Algorithms | Object Oriented Methodologies | Operating System | Database Management System | Machine Learning | Artificial Intelligence | Data Mining | Generative AI | Data Visualization | Cybersecurity

Volunteer Experience

Overall Coordinator - Robita Club (Generative AI), IIIT Allahabad Jan 2023 – June 2025

- Spearheaded a **10+** member team to execute **3+** AI initiatives with PyTorch, orchestrating hands-on workshops.

Achievements

- Secured **99.45 percentile** in JEE Mains 2022, ranking in the top **0.65%** nationwide.
- Achieved Codeforces Specialist (rating 1534) and LeetCode 1900+ in C++.
- Completed Options 101 course by Akuna Capital, mastering foundational options trading concepts [Link - [🔗](#)].