

# Indranil Bain

Portfolio: [not-indro.github.io/Portfolio](https://not-indro.github.io/)

GitHub: [github.com/not-indro](https://github.com/not-indro)

## EDUCATION

### • Indian Institute of Engineering Science & Technology, Shibpur

*B.Tech, Computer Science & Technology, CGPA: 7.83/10*

Email: [indranilbain14@gmail.com](mailto:indranilbain14@gmail.com)

LinkedIn: [linkedin.com/in/indranil-bain](https://linkedin.com/in/indranil-bain)

Howrah, India

December 2020 – July 2024

## SKILLS SUMMARY

- **Programming Languages:** Python, SQL, C++
- **ML Frameworks:** PyTorch, PySpark, TensorFlow, scikit-learn, spaCy, Keras, OpenCV, Transformers
- **LLM & GenAI:** Fine-tuning, RAG, Prompt Engineering, LangChain, AWS Bedrock, Hugging Face, AutoGen
- **Data Pipelines & MLOps:** Apache Airflow, AWS SageMaker, Docker, A/B Testing, ETL/ELT, CI/CD for ML
- **Database & Cloud:** AWS (RDS, S3, Glue, Redshift), Azure, PostgreSQL, MongoDB, FAISS
- **Data & Dev Tools:** Git, Jira, Confluence, DBeaver, Adv. MS Excel, Power BI, VS Code

## EXPERIENCE

### • EvenFlow AI

#### *Data Analyst*

Northfield, IL, USA (Remote)

January 2025 – Present

- Developed a repair-category prediction model for operation codes using TF-IDF and Logistic Regression. Cleaned and prepared the data, tuned the model, and reached 91% accuracy, which helped reduce manual tagging work by 80% and improved repair tracking.
- Created an outreach tool for the support team to easily upload CSV files and trigger customer reminders. Set up a rule-based system plus integrated TextGrid API for scheduled messages, which cut a lot of manual follow-ups (around 75%) and improved rescheduling rates.
- Worked directly with the CEO to dig into operational data, improved the internal calculations, and supported service center managers with insights that helped them streamline processes and drive better results.

### • i3 Digital Health

#### *Data Engineer Trainee*

Bangalore, India

September 2024 – December 2024

- Automated data pipelines to pull drug, cancer, and biomarker data from the ClinicalTrials API using SciSpacy NER and BioBERT model, reduced manual lookup work by 85%. Handled data cleaning, entity mapping, and tagging so the research team could use the data directly.
- Developed a scalable ETL scraper for FDA oncology drug label PDFs, automating ingestion, parsing, and storage (S3) saving 3 Hours/day of manual work. Used BeautifulSoup4 + async task queues.

### • IEST, Shibpur

#### *Summer Project Intern (Faculty Mentor: Prof. Abhik Mukherjee)*

Howrah, India

May 2023 – July 2023

- Developed an GRE practice platform using React. Focused on simplifying the UI and on-boarding flow, which helped students get started faster and led to 35% higher engagement and 40% drop in on-boarding time.
- Implemented a basic recommendation system using collaborative filtering and user behaviour data to suggest practice questions, which helped students improve scores by 25%.

### • IIT Kharagpur

#### *Research Intern (Faculty Mentor: Prof. Debdoot Sheet)*

Kharagpur, India

Jul 2022 – Oct 2022

- Worked on medical imaging pipeline for the Skin Curate app, focusing on cleaning and improving dermatology image data. Improved preprocessing and evaluation steps, helping reduce false positives by 30%.
- Implemented Federated Learning for dermatology image classification, improving model performance and privacy compliance achieving 83.2% accuracy across distributed training nodes.

## PROJECTS

### ◦ Biased News Classifier (Github)

Fine-tuned DistilBERT on the MBIC dataset for bias detection, with preprocessing, class-balancing, and AdamW optimisation. Achieved 73% accuracy and deployed with live NewsAPI ingestion and an interpretability dashboard.

### ◦ pdfSAGE - Chat With PDF (Github)

Built an PDF assistant using LangChain + Pinecone + OpenAI SDK for semantic search & Q/A over documents. Developed Next.js app with PostgreSQL, AWS S3 and chat interface.

### ◦ Handwritten Mathematical Expression Recognition Model (Github)

Contributed to a ViT to GPT2 encoder-decoder for HMER Model, producing LaTeX from images and improving CROHME accuracy to 88% across diverse handwriting styles

## ADDITIONAL PURSUITS

### ◦ REVELATION (Departmental Fest): Assistant Treasurer (2023), Event Secretary (2024)

### ◦ REBECA (Cultural Fest): Publication Co-Lead (2023), Cultural Secretary (2024)

### ◦ SPORTS: Represented our university team in inter-college badminton and football tournaments.