

Ayush Shaw

Durgapur, India | +91-6290591230 | ayushaks099@gmail.com | [LinkedIn](#) | [GitHub](#) | [LeetCode](#)

Education

National Institute of Technology, Durgapur

Aug 2023 – Jun 2027

B.Tech in Computer Science CGPA: **7.91/10** (4 semesters)

Delhi Public School, Durgapur

Graduated: May 2022

Grade: 9/10

Technical Skills

Languages: Python, C++, SQL

Frameworks & Libraries: TensorFlow, PyTorch, Scikit-learn, OpenCV, Flask, Streamlit, LangChain

AI/ML & GenAI: Machine Learning, Deep Learning, NLP, Transformers, RAG, LLMs, GANs, YOLO.

Achievements

- Solved **800+** LeetCode problems; Rated **1851** (Top 4% globally).
- **Specialist** at Codeforces [\[Profile\]](#) & **4 star** coder at CodeChef [\[Profile\]](#).
- Scored **98.1** percentile in JEE Mains (AIR 22K / 1.6M).
- Completed certifications in **Data Science, GenAI, Decode C++** from PW Skills and Udemy.

Experience

Project Intern (Mentored), Techsaksham (Microsoft + SAP)

Jan 2025 – Feb 2025

- Built a multi-disease predictor using **Scikit-learn, Flask, Streamlit**.
- Achieved **88% accuracy** across 4 diseases using cross-validation techniques.

Project Intern, 1Stop.ai

Apr 2025 – Jul 2025

- Created a **hate speech classifier** using **NLTK** and **Decision Tree** with exploratory analysis.
- Built a **heart disease mortality predictor** using **Scikit-learn** and evaluation metrics.

Selected Projects

Agentic RAG — Multi-PDF RAG Chatbot [\[GitHub\]](#)

Python, Streamlit, LangChain, Google Gemini, ChromaDB — Built a production-ready **Retrieval-Augmented Generation (RAG)** platform with **multi-PDF ingestion, multi-user authentication, and persistent storage**. Integrated reranker training, ColPali/ColQwen visual-text retrieval, and streaming LLM responses.

ARGO RAG Explorer — Oceanographic RAG + ML Platform [\[GitHub\]](#)

Python, Streamlit, xarray, SQLAlchemy, LangChain, ChromaDB, scikit-learn, XGBoost, LightGBM, Plotly — Built an **end-to-end RAG-enabled ML platform** for **ARGO ocean data ingestion, visualization, querying, and prediction**. Converted **NetCDF profiles** into a queryable database, integrated **LangChain RAG with Chroma** for **provenance-tracked answers**, and implemented a **multi-model ML workflow** (RF, XGBoost, LightGBM) with **uncertainty estimation**.

AI Meeting Summarizer — Real-Time Meeting Intelligence Platform [\[GitHub\]](#)

Electron, Python, Flask, Socket.IO, SQLite, Deepgram, LLMs — Built a cross-platform desktop application to **record meetings, perform real-time speech-to-text transcription, and generate AI-driven summaries and action items**. Implemented **WebSocket-based streaming, task extraction**, and integrations with **Notion, Jira, and Google Calendar** for workflow automation.

Voice Assistant [\[GitHub\]](#)

Flask, LangChain, Python, SpeechRecognition, VLC, OpenAI/Groq API — Built a voice assistant to **answer queries, play music, set reminders, and fetch weather** using **LLM-based intent detection**. Integrated **YouTube audio playback** via **yt-dlp** and **python-vlc**.

Publications

- Ayush Kumar Shaw, *Transformer-Based Grammatical Error Correction Using Sequence Modeling*, [TechRxiv](#), [\[Code\]](#), July 2025.
- Ayush Kumar Shaw, *Semantic Segmentation for Clothing Detection using Deep Learning*, Preprint at ResearchGate, [\[Code\]](#), July 2025.