

Mayank Joshi

📞 9518267398 📩 joshimayank08012007@gmail.com 🌐 github.com/lothnic 💬 [mayankjoshi0801](https://www.linkedin.com/in/mayankjoshi0801)

Education

Indian Institute Of Information Technology, Una <i>Bachelor of Technology in Electronic And Communication Engineering (CGPA: 8.23)</i>	Expected May 2028
	Una, HP

Experience

- **Indian Institute Of Information Technology, Una - Machine Learning Developer** Oct 2025 - Present
 - Built a machine learning model utilizing genetic algorithms to optimize course scheduling for the college website.
 - Designed and implemented the genetic algorithm in Python to find efficient timetable configurations.
 - Working on integrating the optimization model with the website's backend, ensuring seamless data flow and processing.

Projects

LINGO: Kangri-Hindi NMT for Argos Translate | Python, PyTorch, OpenNMT-py, CTranslate2, SentencePiece, Modal

- Built the first **installable Kangri-Hindi MT package** for Argos Translate, enabling offline translation for a low-resource language spoken by 2M+ people.
- Trained a **6-layer Transformer from scratch** on **26,779 parallel pairs** via end-to-end pipeline (Modal GPU, CTranslate2 INT8); deployed 53 MB model with 150ms CPU latency via CLI/GUI.
- Established the first reproducible Kangri-Hindi NMT benchmark, achieving **BLEU 9.89 (Kangri→Hindi)** and **9.04 (Hindi→Kangri)** with chrF++ 27.50/28.16 on a **2,500-sentence held-out test set**, providing a strong baseline with significant room for improvement.
- Released model files at github.com/lothnic/LINGO; ready for human evaluation and domain adaptation.

FloatChat - LLM-powered oceanographic analysis platform | Python, PostgreSQL, Airflow, RAG, LangChain, FastAPI

- Built an AI-powered data platform to query and analyze global **Argo float NetCDF datasets** containing ocean temperature, salinity, and pressure profiles.
- Processed and filtered **580+ NetCDF float files** using Xarray and stored structured data in a relational database.
- Implemented **RAG (Retrieval-Augmented Generation)** with vector embeddings of database schema and metadata for natural language-to-SQL query generation with a FastAPI backend.
- Integrated a **FastAPI backend** that allows users to ask conversational questions (e.g., temperature trends, salinity anomalies) and receive AI-generated SQL-backed insights.

DeltaVision – Visual Difference Engine | Python, OpenCV, PyTorch, FastAPI, Streamlit, CNN, ConvLSTM

- Built an AI-powered visual inspection system detecting image differences with **90% faster inspection time**.
- Implemented a **convolutional feature extractor** for change localization and a **ConvLSTM temporal module** for tracking gradual degradation.
- Created an interactive Streamlit dashboard with **RLHF-inspired feedback** for accuracy improvement.

Skills

Large-Language-Model Engineering: Transformers, OpenNMT-py, CTranslate2, SentencePiece, KV-Cache, GQA, MoE, HuggingFace, PEFT/LoRA, vLLM, QLoRA, Low-resource NLP

Deep-Learning Stack: PyTorch, TensorFlow, OpenCV, NumPy, SciPy, Pandas, WB, MLFlow, ConvLSTM, Computer Vision

Data Engineering: SQL (PostgreSQL, MySQL), MongoDB, Xarray (NetCDF), Vector Embeddings, RAG, LangChain, Time-Series, Statistical Analysis

Back-End / SWE: Python, C/C++, FastAPI, Flask, RESTful APIs, Docker, Conversational AI, Streamlit

Cloud MLOps: AWS (S3, EC2, Lambda), GCP Vertex AI, Modal, Argos Translate, INT8 Quantization, CI/CD, Git, Linux

Certifications

Data Analyst Associate – DataCamp (*SQL, data wrangling, visualization, business insights*)

Machine Learning Specialization – Andrew Ng (Coursera), DeepLearning.AI (*supervised learning, feature engineering, model evaluation*)