DAKSHBIR SINGH KAPOOR

SOFTWARE, ML-DL, GEN AI, RESEARCH ENGINEER

8130362631 | dakshbir.singh.ug23@nsut.ac.in | Linkedin | GitHub | Leetcode | Codeforces

EDUCATION

B.Tech Computer Science	2023-2027	Netaji Subhas University of Technology, Delhi	**8.73** CGPA
CBSE (Class XII)	2022-2023	K.R.Mangalam World School, Vikaspuri	**97%**
CBSE (Class X)	2020-2021	K.R.Mangalam World School, Vikaspuri	**98%**

WORK EXPERIENCE

Microsoft India SWE Intern, Bengaluru (Upcoming) - Got the offer. Expected May 2026

Google Summer of Code (GSoC) 2025 - Open Climate Fix (June 2025 - August 2025) --Link

- Implemented a high-performance, cloud-native data streaming solution for a **860GB** satellite imagery dataset by integrating Ice-Chunk and Zarr, resulting in a throughput of **31,281.96 MB/s**—a **2.09x** improvement over standard Zarr streaming.
- Authored a comprehensive benchmarking suite to fine-tune data access patterns, including chunking strategies, memory management etc.
- Tech Stack: Ice Chunk, Zarr 3, xarray, Dask, Google Cloud Storage

SCG Foundation (Startup in Stealth Mode)- AI/ML Intern (May 2025)

- Architected **AI-powered career counseling platform** using GPT-4o-mini and ElevenLabs Conversational AI, processing **290+** career profiles with **50+** psychometric reports and **20+** mock interviews.
- Developed intelligent session management with real-time voice processing and a FastAPI backend, achieving **95%** accuracy in career recommendations through advanced prompt engineering and RAG implementation.
- Tech Stack: GPT-4o-mini, ElevenLabs AI, FastAPI, PostgreSQL, Vue.js, OAuth 2.0, RAG, Python

Sabudh Foundation – Data Science Intern (May 2024 - July 2024)

- **Narrative Research on UK General Elections 2024** (Twitter), Scraped and analyzed **65K+** tweets from key UK political figures to predict election outcomes using AI pipeline for sentiment analysis, topic modeling, and NER.
- Identified **10+** critical voter sentiment shifts and **5** high-impact viral tweet patterns influencing election discourse through advanced clustering, comment analysis and EDA.
- Tech Stack: Selenium, NLP, Sentence Transformers, Clustering, EDA

PASSION PROJECTS

DeepSeaAI – AI-powered Marine eDNA Biodiversity Discovery System- an open-set CNN + clustering pipeline classifying 80k+ species from 20M+ eDNA reads and flagging low-confidence sequences as novel-taxa candidates. (Link)
Developed a high-throughput TFRecord-based genomics and 3D analytics workflow cutting compute by ∼60–70% and enabling uncertainty-aware biodiversity mapping. (**Ongoing research paper**)
AI-SQL Query Agent – LLM-based backend for automated SQL generation and correction using Groq API, ew-shot learning, and async processing. (Link**)**
Fitness Chatbot (Gen AI) – RAG-enabled fitness chatbot with LangChain and Llama2 API for health insights.(Link**)**

SKILLS

Languages & Core Technologies: Python, C++ (DSA), NLP and Data Science, ML, DL(CNN, RNN, Object Detection), LLMs, Generative AI
Frameworks & AI Tools: Hugging Face, Github, GPT-4o-mini, ElevenLabs AI, Prompt Engineering, RAG, API's
Core Concepts: OOPS, DBMS, Operating Systems, Data Structure and Algorithms, Computer Networks

POSITIONS OF RESPONSIBILTY

- Placement Cordinator (PC) at **Training and Placement NSUT**.
- Tech and Research Lead at **Enactus NSUT** , Tech Team Member at **IIF NSUT.**

ACHIEVEMENTS

- Selected as 1 of 6 contributors for Google Summer of Code 2025 at Open Climate Fix, from **250+** proposals, with an **acceptance rate** of less than **2.5%**
- JEE ADVANCED **ALL INDIA RANK-6147 and** JEE MAINS **99.37 percentile**
- **10 CGPA Holder throughout my school tenure . Recipient of scholarships, certificates of merit and School-Topper in CBSE for class X and XII.**
- Got the **1st Rank** in the school in "National Science Olympiad" in class XII.
- Leetcode **250+** DSA Problems, Codeforces **1304 rating**
- **Top 100** teams in the Institute in **SIH 2025,** Institute public leaderboard **rank 4** in **Amazon ML Challenge 2025.**