

ABHISHEK SAGAR SANDA

2 Wyman Place, Jamaica Plain, Boston, MA 02130

Email: sabhisheksagar200@gmail.com | Mobile: +1 857-395-9451 | [LinkedIn](#) | [Hugging Face](#) | [Portfolio](#)

OBJECTIVE

AI/NLP engineer with hands-on experience fine-tuning and deploying LLMs for software automation, code generation, test synthesis, and chatbot integration. Skilled in Python, PyTorch, Hugging Face Transformers, LangChain, and embedded systems support. Passionate about building intelligent development assistants that streamline engineering workflows.

EDUCATION

Northeastern University, Boston, MA

Master of Science in Information Systems | GPA: 3.83 | (Expected December 2025)

Relevant Coursework: Theory and Practical Applications of AI Generative Modeling, Advanced LLM Techniques, NLP Engineering

SKILLS

- **Programming:** Python, JavaScript, Java, C, C++, C#
- **Frameworks/Libraries:** PyTorch, TensorFlow, OpenCV, ReactJS, NodeJS, Flask, HuggingFace, Transformers, LangChain
- **ML Focus:** LLM fine-tuning, Transformer models, Reinforcement Learning, Sequential reasoning
- **NLP/LLM Tools:** OpenAI API, LLaMA, GPT-3.5/4, SentencePiece, Codex, BPE
- **Tools:** CUDA, REST APIs, MongoDB, MySQL, ChromaDB, Git, VSCode integration
- **Scientific Domains:** NLP, Computer Vision, Translation Systems, Public Safety
- **Prompt Engineering:** Chain-of-thought prompting, RAG pipelines, few-shot learning
- **Embedded Systems:** Firmware logic structures, hardware interfacing (Raspberry Pi, Arduino, NVIDIA Jetson)

PROFESSIONAL EXPERIENCE

Research Software Engineer Intern, AI for Public Safety | Virtual Presenz, MA | Sept 2024 – Dec 2024

- Fine-tuned YOLOv8 and GPT-4 for CV + NLP multimodal applications (weapon detection + report generation).
- Designed automated labeling tools that integrated natural language prompts with annotated JSON output.
- Built chatbot support system for public safety officers using LLM chaining and context-aware memory.
- Collaborated with cross-functional hardware/software teams to optimize LLM tools for on-device use.

Graduate Engineer Trainee, Full Stack Development | HCL Technologies, Chennai, India | Aug 2022 – Aug 2023

- Delivered secure and scalable .NET-based enterprise applications, achieving 90%+ project efficiency.
- Applied agile methods to reduce support tickets by 10%, boosting user experience through adaptive back-end logic.
- Strengthened application security by resolving vulnerabilities, decreasing SAST/DAST issues by 20%.

PROJECTS

RAG-Powered University Chatbot

- Built an end-to-end AI-powered chatbot system for Northeastern University using Python, FastAPI, Scrapy, and ChromaDB, enabling real-time natural language search and Q&A over hundreds of scraped university web pages.
- Engineered a robust data pipeline for large-scale web scraping, semantic document indexing, and retrieval-augmented generation (RAG), with a modern web frontend and automated data management for production readiness.

Neural Machine Translation for Low-Resource Languages | [Github](#)

- Trained a PyTorch-based Transformer model to translate low-resource language pairs (English–Manipuri) with custom tokenization (BPE + SentencePiece).
- Employed prompt tuning and dataset augmentation to handle ambiguous source structures.

AI-Powered Richard Wyckoff Trading Assistant | [Github](#)

- Developed a GPT-enhanced RL agent for financial market decision-making using Q-learning + custom dataset of 19,000 QA pairs.
- Integrated temporal sequence modeling and NLP to simulate multi-step agent responses in a trading environment.

Dynamic Scientific Chatbot (Roli.ai Hackathon Winner) | [Github](#)

- Designed a GPT-based assistant capable of domain reasoning, retrieval augmentation, and multi-turn scientific interactions.
- Custom-trained LLM prompts to follow lab procedures and react to real-time user intent using OpenAI API.