

SIDDHANT GUPTA

siddhant.gupta1410@gmail.com | [+91-9557575905](tel:+91-9557575905) | [GitHub](#) | [LinkedIn](#) | [LeetCode](#) | [Medium](#)

Education

Graphic Era Hill University, Dehradun, Uttarakhand

June 2025

Bachelor of Technology in Computer Science Engineering, **CGPA: 8.00**

Relevant Coursework: *Software Development, DSA, Machine Learning, DBMS, OS*

Times World School, Dehradun, Uttarakhand

March 2021

12th CBSE, **CGPA: 8.68**

Times World School, Dehradun, Uttarakhand

March 2019

10th CBSE, **CGPA: 9.00**

Internships

Goalwit Technologies - Nikah Forever | SDE Intern | [Link](#)

September 2024 - November 2024

Matrimony Company with over 85,000 successful marriages across India and UAE.

- Devised a matrimonial recommender system using **KD-KNN-LR** algorithm optimizing processing efficiency by **40%**.
- Analyzed **500,000+** profiles using **NumPy**, **Pandas**, **label encoding**, and **Min-Max Scaling** for efficient usage.
- Designed **50+ test cases** and manually tested the user profiles to validate model accuracy and performance.
- Developed **scalable**, modular components and resolved **10+ bugs** through testing and cross-team collaboration.

Projects

AI Career Guidance System Multi AI-Agent | [Link](#)

July 2025

Streamlit, LangChain, RAG, CrewAI, Cloud Deployment, Mistral

- Deployed an **AI** career assistant leveraging **RAG**, **Mistral Large**, **SerpAPI** and **LangChain** to analyze **50+** professions.
- Evaluated 3 LLM providers (OpenAI, Gemini, Mistral) and selected Mistral for production due to generous free tier limits.
- Automated email report generation using **CrewAI** and **SMTP** tools to deliver personalized career reports.

Multi Modal Chatbot (AI Doctor) | [Link](#)

May 2025

LangChain, Groq, Speech Recognition, Generative-AI, Gradio, LLMs

- Coded a **Gen-AI** powered multimodal medical chatbot with speech-to-text, computer vision diagnosis, and text-to-speech response using Python and **GROQ** API.
- Integrated **LLaMA-4** for real-time image analysis and **Whisper-large-v3** for audio transcription with a 20s timeout.
- Deployed an interactive **Gradio** app supporting simultaneous voice and image input.

Image Based Culinary Assistant App | [Link](#)

May 2025

Deep Learning, NLP, LangChain, Python, TensorFlow, Image Processing, RAG, Recommender System

- Developed a **CNN** model with image preprocessing on 5000+ images across 25 categories, achieving **82%** accuracy.
- Built a RAG pipeline for recipe generation using FAISS for retrieval, Sentence-BERT for embeddings, and Gemini API.
- Fine-tuned **Helsinki-NLP English-Hindi** translation model on **10000+** pairs, achieving **BLEU** score of **40.6**.
- Integrated a content-based recommendation engine leveraging **cosine similarity** on 8 culinary attributes.

File Compressor Huffman Encoding | [Link](#)

August 2024

Data Structures & Algorithms, C++, Python, Flask, Cloud Deployment

- Deployed a cloud-based **lossless** file compression tool using **Huffman Encoding** and **Flask**, supporting cross-device compression and decompression.
- Applied **OOPs** principles to develop a space- and time-efficient encoding algorithm, reducing file size by **30%**.
- Leveraged 3 data structures - Binary Trees, Maps, and Heaps, optimizing **runtime ($O(n \log n)$)** and memory usage.

Achievements and Certificates

- IBM Skills Cloud Retrieval Augmented Generation (RAG with LangChain) **Skills:** *RAG Pipeline, Embeddings* | [Link](#)
- Infosys Springboard Agile Scrum Practice **Skills:** *Agile Methodology, Scrum Framework, Sprint Planning, TDD* | [Link](#)
- Solved **270+ LeetCode** problems, strengthening core DSA (graphs, trees, searching) and problem-solving skills. | [Link](#)
- Selected for the pre-final round of Code for Bharat 2.0 Hackathon, ranking among the top **300** teams nationally.
- Published 2 quality datasets on **Hugging face** with over **5000** images and **10000+** parallel sentences.

Skills

Languages: C/C++, Python, Java, SQL, HTML, CSS, JavaScript

Technologies: Machine Learning, Deep Learning, Natural Language Processing, Generative-AI, Recommender Systems

Natural Language Processing: Pipeline, Text-Vectorization, Lemmatization, Stemming, LLMs, NLTK, SpaCy

AI / ML: Regression, Classification, Clustering, Ensemble Methods, NumPy, Pandas, Data Visualization, Neural Networks

Frameworks & Tools: TensorFlow, PyTorch, Scikit-Learn, Flask, LangChain, Git, GitHub, Docker (Basics), IntelliJ (IDE)

Concepts: Data Structures & Algorithms (DSA), Object Oriented Programming (OOPs), SDLC, Operating Systems, DBMS, System Design, Unit-Testing, Code Reviews, Debugging