

SIDDHANT MAHAJAN

+917710983852 | sidumahajan@gmail.com | [LinkedIn](#) | [GitHub](#)

EDUCATION

VIT Bhopal University

(Aug, 2022 – Expected May, 2026)

B. Tech CSE with Specialization in AIML | CGPA: 8.70

SKILLS

Languages: Strong grasp on Python, C++; familiarity with Java, SQL in MySQL.

Experience: Machine learning and Deep learning using Scikit-learn, Pandas, TensorFlow, Rapids; Basic Data Analytics using MS Excel, Dataiku, Tableau.

PROJECTS

DocuSummary

(July, 2025)

Python, Transformers, VLMs, Generative AI

- A RAG application for streamlining Data Extraction and contextual Query answering using LLMs for PDF documents.
- Utilized Qwen 2.5 VLM, Qwen 3 Embeddings Model and OpenGVLab's InternVL3 model for extraction of textual data description, paired with ChromaDB for fast and scalable similarity searching for related contextual chunks.
- Optimizations to speed and reduced hardware dependency by utilizing quantized models for RAG Pipeline.

Brain Cancer Detection using Fine-tuned CNNs

(May, 2025)

Python, TensorFlow

- Designed Functional Ensemble models, utilizing MobileNetV3, VGG19 and ResNet50V2 models for finetuning.
- Achieved 96% accuracy on unseen test data of 1000 images.
- Reduced the time taken for training and predictions with the functional model architecture over traditional Ensemble architecture, allowing for expansion of model scopes in future with modular design.

Lungs Disease Detection through Audio Recordings

(April, 2025)

Python, TensorFlow, Librosa for Audio

- Created a CNN-RNN model to accurately map and identify 8 classes of diseases using audio samples.
- Extraction of audio features using MFCC coefficients, allowing for feature extraction using single dimension CNN filter. The output is then passed as a time series data to RNN layers, providing robust pattern identification and recognition capabilities.

CO – CURRICULARS AND EXTRA – CURRICULARS

CO – CURRICULARS

Sabudh Foundation – STPI, India

(January 2025 – June, 2025)

- Completed Data Science Internship at Sabudh Foundation (STPI Initiative), gaining hands-on experience with machine learning, deep learning, and natural language processing to solve real-world data problems.
- Utilized Dataiku for data wrangling, model training, and deployment, enhancing efficiency and collaboration in data science workflows.

Certificates

- Machine Learning with TensorFlow – FreeCodeCamp
- Applied Machine Learning in Python – Coursera
- Dataiku Developer Certificates – Dataiku

EXTRA – CURRICULARS

National Service Scheme - VIT Bhopal University

- Spearheaded planning for women's empowerment events during 'International Women's Day', resulting in increased local participation and heightened awareness of gender equality issues among villagers.
- Part of NSS – VIT Bhopal 10,000 trees plantation drive, managing a team of 10 volunteers along with coordinating with other 25 teams to make the drive successful.

ADDITIONAL INFORMATION

Languages: English (Professional working proficiency), Hindi and Marathi (Native professional proficiency)

CODING PROFILE

[GeeksForGeeks](#)

[Leetcode](#)

[ML Notebooks - GitHub](#)