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 (June, 2025)

- A RAG project that utilizes LLMs for embedding, searching and relaying answers to questions based on an external knowledge source. Built using Gradio for chat interface.
- Utilizing IBM's Granite Models and Embeddings, documents can be converted into knowledge vectors, which are stored into a database.
- The database is then utilized by the LLM model to answer certain queries, thus ensuring authenticity of the knowledge along with high quality outputs of LLMs.
- Ensures all data can be processed locally, thus preventing leakage of sensitive data and preserving confidentiality.

Brain Cancer Detection using Fine-tuned CNNs (May, 2025)

- 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, as tested locally

Lungs Disease Detection through Audio Recordings (Apr, 2025)

- Created a CNN-RNN algorithm pipeline to accurately map and identify 8 classes of patient diagnoses using audio samples.
- Extraction of data using MFCC coefficients, thus converting audio into spatial frequency domain; allowing for feature extraction using single dimension CNN filter. The output can be passed as a time series data to RNN layers, providing robust pattern identification and recognition capabilities.
- Trained on over 870 audio samples, the model achieves exceptional accuracy of 94.8% on a testing set of 97 data samples, thus proving its feasibility on aiding for a better and robust medical usability.

CO-CURRICULARS AND EXTRA-CURRICULARS

CO-CURRICULARS

- HackVega 2025 Semi-finalist, demonstrating the grit to learn and compete not only in terms of coding but also on logical and aptitude basis.
- Zelestra x AWS ML Challenge – positioned within top 150 teams, showcasing deeper understanding of concepts such as Data Analytics, feature engineering, model feature selection and heuristic understanding.

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.

Sabudh Foundation – STPI, India

- Selected as a Data Science Intern in the Sabudh Foundation Initiative, backed by STPI, Ministry of Electronics & Information Technology Government of India.

ADDITIONAL INFORMATION

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

CODING PROFILE

GeeksForGeeks Leetcode ML Notebooks - GitHub