GYANBARDHAN

6204487848 bt21cse194@iiitn.ac.in gyanbardhan Gyanbardhan

EDUCATION

Indian Institute of Information Technology, Nagpur **Graduating June 2025**
Bachelor of Technology in Computer Science *7.59 CGPA*
Relevant coursework : Deep Learning,Artificial Intelligence, Machine Learning, DBMS,Data Structures and Algorithms,OOPs

SKILLS

Languages Python, SQL, C, C++, Java
Technology Generative AI, Machine Learning, Deep Learning, Artificial Intelligence, Natural Language Processing (NLP), Retrieval Augmented Generation (RAG), Large Language Models (LLM), Data Manipulation, Data Analysis, Statistical Analysis, Data Collection, Data Preprocessing
Frameworks TensorFlow, Scikit-Learn, Keras, Hugging Face, Langchain, Streamlit
Databases MySQL, PineCone, Chroma DB
Tools Matplotlib, Seaborn, Jupyter Notebook, Google Colab, VS Code, Kaggle, Microsoft Office, GIT

PROJECTS

Gemini Student ***- Github - Web Link - Demo*** **June 2024 – July 2024**
*Created an educational web application having **5** interactive components using the **Gemini large language model** for students.*
- Developed **3** key components: **Chatbot** for real-time responses, **Image QA Gemini** for image-based Q/A, and **QA Gemini** for insightful question-answering.
- Carried out **MCQ GEN** to generate PDF with multiple-choice questions and verified answers based on **4** user inputs, including **text, topic, number of questions**, and **difficulty level**
- Implemented a **RAG** application, **Chat with PDF** for efficient document retrieval and QA with uploaded PDFs using the **Pinecone** Vector Database with **768** dimensional **embeddings**.
- Deployed the application on **Hugging Face Spaces** on **16 GB RAM CPU** using **Streamlit** for easy access and user-friendly interaction.

Duplicate Question Detection ***- Github - Demo*** **Apr 2024 – May 2024**
*Designed a **natural language processing (NLP)** system using **streamlit** to identify duplicate questions on a Q/A platform.*
- Applied ***text preprocessing***, techniques like *tokenization, lowercasing, stop words removal, stemming/lemmatization*, and *special character removal* to get accuracy of **80.6** also utilized **Bag of Words (BoW)** and **Term Frequency-Inverse Document Frequency (TF-IDF)** to represent text data.
- Engineered features basic features (e.g., question length, word counts), advanced token features (e.g., common word ratios), length-based features, and fuzzy matching features to improve model performance by **1.2**.
- Achieved final accuracy of **81.77** in identifying duplicate questions with **streamlit deployment**.

Plant Disease Detection Generalization ***- Github - Demo*** **Feb 2024 – March 2024**
Innovated a deep learning-based system for early detection of plant diseases using leaf images
- Put into practice various state-of-the-art **CNN** architectures **AlexNet, VGG-16, VGG-19, ResNet, DenseNet, EfficientNet**, and **ConvNextLarge** and got accuracy in range **96.4 to 99.8**, by training on more than **25k images**.
- Formulated **bagging** technique to improve accuracy by combining outputs of best 3 models.
- Enhanced model **accuracy by 4-5** through the use of *regularizations, dropouts, normalizations, equalizations, clustering, deployment, visualizations, and image segmentation techniques*.

MediChat ***- Github - Demo*** **May 2024 – June 2024**
Built an advanced medical chatbot to assist with clinical queries and provide information based on medical literature.
- Utilized the **Llama-2-7B-Chat model** and integrated Chroma DB for efficient data retrieval, enhancing Retrieval Augmented Generation (RAG) application, response accuracy by **90**.
- Implemented high-quality embeddings using **sentence-transformers/all-MiniLM-L6-v2** to generate **768**-dimensional vectors, improving query handling by **50**.
- Constructed the application in Flask, enabling **24/7** access to medical information with real-time query resolution.

ACHIEVEMENTS AND CERTIFICATIONS

- Secured 8th rank as a finalist in **NextGenEd**: *Crafting the Future of Learning* at IIITDM, Kurnool.
- Scored **19th** rank in **Datapunk, Tecnovate'24 IIIT Naya Raipur** to build a ML Model
- Mastered **Generative AI** and Large Language Models, Scored 82.35 in quizes and 100 in assignments.
- Completed successfully Stanford University's Coursera course on **statistics** equipped me with essential skills for data analysis, enhancing my statistical proficiency
- **Leetcode**- 1701 ratings , 400+ questions on DSA and SQL
- **Codechef**- 3 star , 1620 rating