

DEYYALA SAI VENKAT

📞 8309950589 ✉ venkatd@iisc.ac.in
✉ deyyalasaivenkat@gmail.com [in](#) [Sai Venkat](#)

Education

M.Tech in Computational and Data Sciences

Indian Institute of Science, Bangalore

Aug. 2023 – Current

CGPA: 7.9/10

B.Tech in Biotechnology

National Institute of Technology, Tadepalligudem, Andhra Pradesh

Aug. 2019 – May 2023

CGPA: 9.09/10

Projects

Phenotype Normalization for Rare Diseases (M.Tech)

- Built a phenotype normalization model using **GPT embeddings** (text-embedding-3-small) to map Human Phenotype Ontology (HPO) terms into a vector database for efficient retrieval.
- Applied NER to extract phenotypes from clinical text, using a sentence transformer and **FAISS** for similarity search, facilitating accurate retrieval of the closest HPO IDs for phenotypic terms.

BIOMed_NER: Named Entity Recognition for Biomedical Entities

- Developed a **Named Entity Recognition (NER)** model using **DeBERTaV3** base to tag 41 biomedical entities.
- Utilized the **MACCROBAT2020** dataset with 400 clinical notes, split 80% for training and 20% for validation, to accurately tag biomedical terms in clinical text.
- Achieved high performance in tagging specific biomedical entities, enhancing information retrieval and structuring of clinical data for downstream applications.
- Published the model on Hugging Face as **BIOMed_NER**, making it available for the research and healthcare community.

Sophisticated Medical Chatbot with RAG and LLM Integration

- Developed an advanced medical chatbot using **Retrieval-Augmented Generation (RAG)** and **Llama 2** as the LLM, integrating **Sentence Transformers** for embedding and **FAISS** for efficient vector retrieval.
- Embedded critical medical knowledge from *Harrison's Principles of Internal Medicine* and *Davidson's Principles of Medicine*, enabling accurate responses to complex healthcare queries.
- Implemented the chatbot using the **Chainlit** library for an interactive conversational interface, allowing seamless AI-driven assistance in healthcare scenarios.

GPT-2 Implementation for Text Auto-Completion Tasks

- Implemented and trained a GPT-2 model from scratch using pytorch on the **FineWeb-Edu** dataset of 10 billion tokens. Validated the model using the **HellaSwag** dataset, achieving an accuracy of 31.33%.
- Utilized Distributed Data-Parallel (DDP) to efficiently leverage 2 RTX A6000 GPUs (48GB each) for training.

Machine Translation using Transformers

- Built a **Transformer model** from scratch using PyTorch for English-Kannada translation on a dataset of 300,000 sentence pairs.
- Implemented **subword tokenization** and optimized training with **gradient accumulation** and **learning rate scheduling**.
- Achieved a **BLEU score of 0.26**, demonstrating effective translation for a low-resource language pair.

Course Assignments and Mini-Projects

- **Generative Adversarial Networks (GANs):** Developed and trained **WGAN**, **CGAN**, and **DCGAN** architectures on datasets with 90 animal classes, CelebA dataset, and butterfly images, achieving realistic image synthesis.
- **Variational Autoencoders (VAEs):** Implemented **VAEs** and **VQ-VAEs** on butterfly and animal datasets to perform generative tasks, capturing meaningful latent representations.
- **Diffusion Models:** Applied advanced diffusion techniques, including **DDPM** and **DDIM**, to generate high-quality images, gaining hands-on experience with state-of-the-art generative models.
- **Classification Models:** Implemented logistic regression and K-Nearest Neighbors (KNN) for binary and multi-class classification tasks.
- **Neural Networks:** Built Multi-Layer Perceptron (MLP) and Convolutional Neural Networks (CNN) from scratch; tested models on MNIST and CIFAR-10 datasets.

Relevant Coursework

- | | | |
|-------------------------------------|-------------------------------------|---|
| • Machine Learning for Data Science | • Deep Learning for Computer Vision | • Advanced Deep Representation Learning |
|-------------------------------------|-------------------------------------|---|

Technical Skills

Languages & Tools: Python, C++, PyTorch, Matplotlib, Pandas
Technologies: RAG, LLMs, NLP, Computer Vision, Deep Learning

Academic Accomplishments & Extracurricular

- Received Gold Medal for academic excellence at NIT Andhra Pradesh
- **AIR 13** in GATE 2023, Biotechnology
- **AIR 42** in GATE 2022, Biotechnology
- LeetCode Profile: [Sai Venkat](#)