Shubham Mishra

mishra4475@gmail.com | Phone: (+91) 8770867683 | skmishra.netlify.app
GitHub | LinkedIn | Medium

Education

Rajiv Gandhi Proudyogik Vishwavidyalaya — Bhopal, India
B.Tech (CSE) — 2021-2025

Skills

Languages/Frameworks - C/C++ | C# | Python | PyTorch | Tensorflow | Langchain | LlamaIndex | JavaScript | HTML/CSS | SQL | Spacy Librosa | FastAPI | vLLM | Selenium | Streamlit/Gradio | Git | Docker | Kubernetes | GCP | **Domains Explored** – Computer Vision, NLP, Unity

Experience

Artificial Intelligence Intern — **Bangalore, India**

Wysa – Touchkin Pvt. Ltd. — **Aug 2024 – Dec 2024**

- Optimizing models for faster inference and performance.
- Migrating Wysa's internal NLP-AI tools from a Flask to a FastAPI-based implementation.
- Training and fine-tuning language models to boost the NLP capabilities of the Wysa mental health app directly enhancing the experience of over a million active users.

Deep Learning Intern (R&D) — **Delhi, India**

DeepLogic AI — **March 2024 – May 2024**

- Worked under the R&D team and engineered various high-throughput retrieval pipelines to scale and replace, Vectara endpoints within DeepLogicAI's enterprise search solutions.
- Conducted rigorous trials across various versions of RAG pipelines for each component such as re-rankers, chunking, and indexing, and finetuning multilingual embedding models on custom datasets.
- Our retrieval system outperformed high-precision retrieval models such as ColBERT (v2).

AI Research Intern — **Remote**

FireLLama Pvt. Ltd. — **Feb 2024 – May 2024**

- Explored various Vision-Language Models (VLMs) to replace the PaddleOCR solution, enhancing the retention of document structures. Managed the entire process from *Proof of Concept* to deployment.
- Developed Speech APIs for chat-bots leveraging various open-source speech models.
- Fine-tuned VLMs and used pre-trained CNNs to create rich embeddings for visual-image search.
- Developed multiple APIs encapsulating different anomaly detection models and writing lengths of unit tests, ensuring smooth integration into the production environment.
- Created custom metrics to evaluate models for classification and Name Entity Recognition (NER) tasks.

Projects

Graph-Enhanced Visual Language Processing – Intersecting Language and Graphs: Python, PyTorch, VLMs, Graphs

- Graph Vision is a Python library that aims to create topological maps for an environment connecting neighboring image segments, capturing each segment's spatial and semantic feature embeddings.
- Enables localization of objects relative to one another in the topology with language description of the objects using Dijkstra's algorithm.

Brain MRI Segmentation for Tumor Detection: Deep Learning, PyTorch, Docker, Streamlit

- Model trained on a diverse dataset encompassing various tumor and non-tumor scans. Capturing the inherent heterogeneity of brain tumors encountered in clinical practice, re-producing the U-Net paper 2015.
- The project has a docker image available on Docker Hub. A user-friendly Streamlit front-end interface on Hugging face Spaces for real-world clinical inference achieving a high validation *Dice score* of *~0.9*.

Pool of Models: PyTorch, ViTs, CNNs

- A GitHub repository containing a variety of Deep Learning architectures implemented from scratch with PyTorch, features both supervised and unsupervised learning models.
- The architecture primarily includes various important ViTs (Swin, Dino, MAE, CvT, etc.). A detailed walkthrough of some of these papers is on my medium page.

Sentence-Level Lipreading: PyTorch, Deep Learning, LSTM

- A Recurrent Network model to predict the spoken sentence by extracting features from the lip movements in the frames.
- Improves upon the paper Based on End-to-End Sentence-level Lipreading by replacing the GRU-based Implementation with a bi-directional LSTM using a CTC Loss to handle the variable length of input alignments (spoken sentence) with Kaiming Normal (He) initialization.

Other Achievements and Contributions

- **Finalist** Inter College Bhopal Hackathon for Deep Learning-based Lipreading
- **5* Star** gold **SQL** on Hackerank
- Partner writer of TheDeepHub publication explaining various deep learning architectures and giving detailed paper walkthroughs.
- Contributed to the winning AI Team in the **Wysa** internal Hackathon.
- **Bronze Medalist** on Kaggle Dataset for dynamic IMDB dataset generation.

Courses

- Deep Learning Specialization by Andrew Ng – DeepLearning.AI — Coursera
- Structuring Database and Management systems with MySQL from Meta — Coursera
- Modern Computer Vision PyTorch, Tensorflow2, OpenCV — Udemy
- Algorithmic Toolbox by University of California San Diego — Coursera