Dristanta Das

LinkedIn: dasdristanta13 Kaggle: dristantadas
dasdristanta13@gmail.com GitHub: dasdristanta13

Skills

- **Expert**: Python (Pandas), R, Git, VSCode, Data Analysis,Visualization
- **Proficient:** Machine Learning, Deep Learning, NLP, NER, Pytorch, Language Model
- **Novice:** Computer Vision, MySQL, LLM, LORA

Experience

Abzooba Inc **Kolkata, West Bengal**
Associate Data Scientist July 2022–Present

- Contributed to building a Provider search system utilizing advanced NLP methods like semantic search, NER, autocorrect, and autocompletion to provide personalized provider suggestions from user queries.
- Leveraged open-source Large Language Models (LLMs) to expedite data preparation, enhancing efficiency by 30-40% while meticulously aligning customer-provided tabular data.
- Implemented semantic search capabilities enabling the retrieval of user-friendly medical terms from unknown layperson terms, culminating in a 20-25% enhancement in the overall search experience through improved query interpretation and accessibility.
- Applied NER techniques, resulting in a significant 15-20% precision boost in search results by extracting key features from user queries, aligning the system with customer needs.

Videonetics Technology Private Limited **Kolkata, West Bengal**
Data Science Intern Jan 2022 — June 2022

- Independently developed an end-to-end Automatic Number Plate Recognition (ANPR) solution employing YOLO-v6 and ResNet architectures in PyTorch. Overcame challenges like low-res license plate images from roadside CCTV cameras.
- Spearheaded solo efforts in data pre-processing and annotation, refining unannotated number plate images using advanced techniques. Enhanced character visibility, yielding a dataset that improved ANPR accuracy by 20%.
- Applied PyTorch and state-of-the-art computer vision models (YOLO-v6 and ResNet) for character segmentation and classification. Achieved a performance boost of 15-20% by meticulously fine-tuning model parameters.
- Crafted a user-friendly UI with Streamlit in Python, seamlessly integrating the ANPR system. Empowered users to upload and process number plate images, showcasing the system's capabilities through intuitive visualizations and displays.

Indian Institute of Technology Kharagpur **Kharagpur West Bengal**
Summer Intern Aug 2021 — Jan 2022

- Independently crafted a deep learning solution using a pre-trained ResNet-34 model, fine-tuning it to analyze breathing sounds, achieving an impressive accuracy range of 70-75% for distinguishing COVID-19 positive and negative patients based on cough sounds despite limited prior work in this area.
- Engaged in the curation and processing of a comprehensive dataset comprising around 1,500 breathing sounds encompassing recordings from COVID-19 afflicted individuals and healthy participants worldwide.
- Spearheaded a pioneering approach by combining hand-crafted features and MFCCs, resulting in a significant enhancement of the deep learning model's performance by 35-40%
- Achieved robust and dependable results by generating probabilities for each sample, providing a measure of the likelihood of an individual being COVID-19 positive or negative, thereby contributing to the advancement of non-invasive COVID-19 detection methodologies.

Education

RKMVERI **Belur, West Bengal**
Big Data Analytics M.Sc. 7.71 CGPA 2020–2022
Presidency University **Kolkata, West Bengal**
Mathematics B.Sc. 7.05 CGPA 2017–2020

🏆 Recognitions

Kaggle Competition "CommonLit - Evaluate Student Summaries", **Top 9%(Bronze Rank)** Nov 2023
Kaggle Competition "Predict Health Outcomes of Horses"; **Top 30%** Nov 2023