Dristanta Das

Kolkata,West Bengal, Phone:+91-8583918089
www.linkedin.com/in/dasdristanta13/ Kaggle: dristantadas
dasdristanta13@gmail.com GitHub: dasdristanta13

Summary

Associate data scientist with expertise in computer vision, natural language processing (NLP), deep learning, and machine learning. Constructed cutting-edge NLP solutions for a healthcare provider search system that increased operational efficiency by 20–25%. Improvement of 30–40% efficiency was gained by using open-source large language models to speed up data preparation activities.

Skills

Expert: Machine Learning, Deep Learning, LightGBM, XGBoost, NLP, NER, PyTorch, Language Model

Proficient: Python, R, Git, Github, Gitlab, Bitbucket, VS Code, Scikit-learn, Data Analysis

Novice: Computer Vision, MySQL, LLM, LORA, AWS(S3, EC2), Azure

Experience

UST (Formerly Abzooba Inc) **Kolkata, West Bengal**
Associate II Data Scientist July 2022 – Present

- Orchestrated a client-facing provider search system through cross-functional collaboration, leveraging advanced NLP techniques to deliver personalized provider suggestions, spearheading optimization endeavors that boosted search efficiency by 30% while exceeding client expectations via clear communication.
- Leveraged open-source Large Language Models (LLMs) to hasten data readying tasks, enhancing efficiency by 40% while using tabular data.
- Collaborated with cross-functional teams to implement semantic search capabilities, enabling the retrieval of user-friendly medical terms from unknown layperson terms, culminating in a 25% enhancement in operational efficiency and ultimately improving service quality for end-users
- Applied NER techniques to product, resulting in significant 15% precision boost in search results by extracting key features from user queries, aligning system with customer needs.

Videonetics Technology Private Limited **Kolkata, West Bengal**
Data Science Intern Jan 2022 – Jun 2022

- Pioneered 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 which in-turn improved performance upto 15%.
- Spearheaded solo efforts for data pre-processing and annotation, refining unannotated number plate images using advanced techniques. Enhanced character visibility, yielding dataset that improved ANPR accuracy by 20%.
- Applied PyTorch and SOTA computer vision models (YOLO-v6 and ResNet) for character segmentation and classification. Achieved performance boost of 15% by fine-tuning model parameters.

Indian Institute of Technology Kharagpur **Kharagpur West Bengal**
Summer Intern Aug 2021 – Dec 2021

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

Personal Projects

Automated Summary Evaluation Competition

Organized by CommonLit, Vanderbilt University, Georgia State University & Kaggle

- Architected a machine learning model that gained a 0.48 Mean Columnar Root Mean Squared Error, ranking in the top 9% of participants.
- Bagged a Bronze Medal from Kaggle for the model's performance in evaluating the quality of student-written summaries.
- Showcased the ability to create practical impact by enabling automated scoring and feedback for teachers and learning platforms.

Education

RKMVERI — **Belur, West Bengal**
Big Data Analytics, M.Sc., 7.71 CGPA — July 2020 – June 2022

Presidency University — **Kolkata, West Bengal**
Mathematics, B.Sc., 7.05 CGPA — May 2017 – June 2020

Certificate & Recognitions

AWS Cerified Cloud Practitioner: CLF C01

Microsoft Certified: Azure Fundamentals, AZ-900