


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 natural language processing (NLP), deep learning, and machine learning. Engineered cutting-edge NLP solutions for a healthcare provider search system, increasing operational efficiency by 20–25%. Catalyzed a 30–40% improvement in efficiency by utilizing open-source large language models to expedite data preparation activities.

Skills

Expert: NLP, NER, Machine Learning, Deep Learning, XGBoost, PyTorch, Statistical modeling

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

Novice: Predictive model, BigQuery, SQL, LLM, AWS, Azure

Experience

UST

Kolkata, West Bengal

Associate III Data Scientist

Oct 2022 – Present

- Orchestrated the development of a client-facing provider search system using advanced NLP techniques, boosting search efficiency by 30% and exceeding client expectations.
- Accelerated data preparation tasks using open-source Large Language Models (LLMs), enhancing efficiency by 40% with tabular data.
- Collaborated with cross-functional teams to implement semantic search capabilities, enabling the retrieval of medical terms from unknown layperson terms, culminating in a 25% enhancement in operational efficiency and ultimately improving service quality for end-users

Associate II Data Scientist

Jul 2022 – Sept 2022

- Spearheaded a comprehensive analysis of an adherence program targeting Eliquis patients using BigQuery, enabling the client, a leading provider of pharmacy-driven patient engagement solutions.
- Engineered comprehensive data transformations and intricate joins to construct a robust dataset; utilized statistical analysis to highlight critical factors, enhancing the client's understanding of adherence drivers by 45%
- Demonstrated the program's impact, indicating an ROI of 800% for the program.

Personal Projects

Retrieval-Augmented Generation (RAG) QA Bot Development

- Engineered a Python-based QA bot leveraging Retrieval-Augmented Generation (RAG) to deliver accurate answers from PDF content, significantly enhancing information retrieval processes.
- Implemented advanced LLMs, including OpenAI's GPT-3.5-turbo, hybrid retrieval methods, and contextual compression to optimize response quality.
- Optimized system performance with caching strategies using SQLite and integrated Slack for real-time result dissemination.

Automated Summary Evaluation Competition

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

- Developed a machine learning model that gained a 0.48 Mean Columnar Root Mean Squared Error, ranking in the top 9% of participants.
- Achieved a Bronze Medal on Kaggle by developing a solution for evaluating student-written summaries, enhancing evaluation accuracy by 65% and outperforming 90% of competing solutions.
- Revolutionized the education platform by implementing automated scoring and feedback systems, reducing teacher manual grading time by 40% and enhancing student learning outcomes by 25%.

Education

RKMVERI

Big Data Analytics, M.Sc., 7.71 CGPA

Presidency University

Mathematics, B.Sc., 7.05 CGPA

Belur, West Bengal

July 2020 – Jun 2022

Kolkata, West Bengal

May 2017 – Jun 2020

Certificate & Achievements

AWS Certified Cloud Practitioner: CLF C01

Microsoft Certified: Azure Fundamentals, AZ-900