Abhijeet Lokhande

Data Scientist & Machine Learning Engineer

07480801382 ⋄ London, United Kingdom
abhijeetlokhande1996@gmail.com ⋄ https://www.linkedin.com/in/ablds/

SUMMARY

I am a highly skilled **Data Scientist & Machine Learning Engineer** proficient in designing and implementing machine learning models to address complex real-world problems. With a strong mathematical background and a passion for innovation, I excel in creating models that provide actionable insights and drive strategic decisions.

TECHNICAL SKILLS

Proficient In	Machine/Deep Learning, NLP, LLMs, Python, SQL, Django Retrieval Augmented Generation(RAG), Data Analysis Scikit-Learn, TensorFlow, Keras, PyTorch, XGBoost ETL processes, Data pipelines, MySQL, PostgreSQL, MongoDB AWS, Git, GitHub, Docker Numerical Optimization, Statistics, Forecasting

PROFILE LINKS

Linkedin	https://www.linkedin.com/in/ablds/
Gitlab	https://gitlab.com/abhijeetlokhande1996
Github	https://github.com/abhijeetlokhande1996
Hugging Face	https://huggingface.co/ab-ai

OPEN SOURCE CONTRIBUTIONS

- Implemented a project inspired by OpenAI's CLIP, developing a **Natural Language Image Search (NLIS)** system that enables efficient image search and retrieval based on textual descriptions using advanced deep learning techniques.
  `https://github.com/abhijeetlokhande1996/text-to-image-search`
- **Fine-tuned GPT(LLM) Model** using LORA to detect Personally Identifiable Information(PII).
  `https://huggingface.co/ab-ai/PII-Model-Phi3-Mini`
- **Fine-tuned Bert Model** to detect Personally identifiable information(**22,000** Downloads on HuggingFace). `https://huggingface.co/ab-ai/pii_model`

WORK EXPERIENCE

Built AI (FinTech & AI Company) April-2022 - Present
Data Scientist *London, UK*

- Achieved 30% improvement in link rate of London property databases, enhancing data accuracy and efficiency, by developing a Token Classification NLP model to detect components(Postcode, Building Name/Number etc) in UK addresses and implementing robust data preprocessing pipeline.
- Developed a Conversational AI Bot for investment brochure analysis, leveraging Large Language Models (LLMs), Vector Databases, Embedding techniques, RAG and React, enabling users to upload multiple brochures and query the system for relevant information retrieval and natural language responses.

- Achieved 60% reduction in rent per square foot error by developing 3D cosine kernel to enhance existing machine learning algorithm, and implementing Leave One Out strategy for error evaluation.
- Reduced project creation time by 70% by improving the Natural Language Processing (NLP) model by integrating AWS Textract for extracting text data from documents, and implementing a custom algorithm that combined Textract's output with a Large Language Model (LLM) for enhanced text understanding, streamlining workflow.
- Boosted performance and responsiveness of financial modeling engine by 70% reduction in execution time, achieved through optimizing existing Python code by implementing Cythonization, Vectorization, and Distributed Computing techniques.

King's College London March 2021 - January 2022
Research Associate *London, UK*

- Developed a novel Deep Learning model utilizing Graph Convolution Network and Generative Adversarial Network (GAN) techniques for Drug Discovery, resulting in a 63% enhancement in the novelty score of drug structures.
- Designed and implemented a robust data preprocessing pipeline for clinical data, ensuring accuracy and completeness of the million rows dataset.
- Utilized Scrapy and Beautiful Soup to gather data from various websites for research purposes

EDUCATION

King's College of London
MSc in Data Science *London, UK*

University of Pune
B.E. in Computer *Pune, India*