

Abhijeet Lokhande

Data Scientist & Machine Learning Engineer

📞 +44 7480801382 abhijeetlokhande1996@gmail.com 🏠 London, UK
🌐 [in/ablds](#) 🐙 [github.com/abhijeetlokhande1996](#)

Summary

Experienced Software Engineer, Data Scientist & Machine Learning Engineer skilled in developing innovative AI solutions, improving data accuracy and efficiency using LLM, Generative AI, NLP, Geospatial Data Science, Machine Learning and Deep Learning techniques. Strong Python development background and solid foundation in data science from King's College London. Experienced in enhancing workflow efficiency through algorithmic improvements and optimization strategies.

Work Experience

Data Scientist & Machine Learning Engineer (*Built AI*) **London, UK** 04/2022 - present

- **Natural Language Processing (NLP):** 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.
- **Data Accuracy & Efficiency:** 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.
- **Geospatial Data Science:** Achieved 60% reduction in rent per square foot error and robust performance improvements by developing 3D cosine kernel that utilized time dimension in time series data to enhance existing machine learning algorithm, and implementing Leave One Out strategy for error evaluation.(Geospatial Data Science)
- **Performance Optimization:** 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.
- **Algorithm Optimization:** Reduced location intelligence map generation time by 40% through optimizing existing algorithm to efficiently handle large volume data.
- **Web-based Q&A RAG Application:** Developed a web-based Q&A application(RAG) for investment brochure analysis, leveraging Large Language Models (LLMs), Vector Databases, Embedding techniques, and Streamlit, enabling users to upload multiple brochures and query the system for relevant information retrieval and natural language responses.

Research Associate (*King's College London*) **London, UK** 04/2021 - 01/2022

- **Deep Learning for Drug Discovery:** 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.
- **Clinical Data Management:** Designed and implemented a robust data preprocessing pipeline for clinical data, ensuring accuracy and completeness of the large volume of data.

Python Developer - Data Science (*Aspect Ratio*) **Pune, India** 02/2018 - 06/2020

- **Forecasting Models:** Developed web-based forecasting model for more than 100 countries to predict patient progression and revenue of a drug.
- **Data Visualization:** Developed a dashboard of polished visualizations to share results of data analyses.

Open Source Contribution

- Fine-tuned Bert(LLM) to detect Personally identifiable information(17,000 Downloads on HuggingFace) ([PII Model](#))

Skills

- LLMs, Name Entity Recognition(NER), Natural Language Processing(NLP), Retrieval Augmented Generation(RAG)
- Generative AI, Deep Learning, Geospatial Data Science
- PyTorch, TensorFlow, LangChain, Django, Flask, FastAPI, Numpy, Pandas, SciKit-Learn, Ray Serve
- Python, C++, R

- MySQL, Postgres, MongoDB, PySpark
- Predictive Modelling, Financial Modeling
- AWS, SQS, Sagemaker, GCP(Vertex AI), Conversational AI
- Docker, MLOPS, Git, Github Actions, CircleCI(CI/CD)
- Numerical Optimization, Distributed Computing

Key Achievements

- Improved project creation times, data accuracy, and model performance through innovative applications of machine learning and optimization techniques.
- Developed advanced models and algorithms that significantly enhanced predictive accuracy and operational efficiency.
- Delivered impactful data-driven solutions across various domains including real estate, finance, and healthcare.

Education

MSc Data Science <i>King's College London</i>	London, UK
Relevant Courses: Machine Learning, Deep Learning, Data Analysis, NLP, Predictive Modelling.	
Computer Engineering <i>University of Pune</i>	Pune, India
Relevant Courses: Software Development, Python, Web Development, Microprocessor Architecture	