

Kamesh Kotwani | Data Scientist

United Kingdom | kameshkotwani@gmail.com | +447932851567 | LinkedIn | GitHub

Summary

Microsoft Certified: Azure Data Scientist Associate (DP-100) with 2+ years of experience building AI models and deploying MLOps pipelines. Proficient in PyTorch, TensorFlow, and cloud platforms (AWS, Azure) with expertise in model tracking (MLflow) and versioning (DVC). Skilled in delivering scalable machine learning solutions and optimizing workflows for production environments.

Skills

Languages: Python, Scala, R, SQL, Bash

Cloud: AWS (SageMaker, EC2), Azure(AZ ML), Databricks, Docker, Kubernetes, CI/CD, MLflow, Ansible, Airflow

Data Management & Big Data Tools: MongoDB, Spark, PySpark, Hadoop, Elastic Search, Kafka

Data Visualization & Analytics Tools: Power BI, Tableau, KNIME Analytics, IBM SPSS

Machine Learning: PyTorch, Keras, TensorFlow, Scikit-learn, XGBoost, Gradient Boosting, Predictive Modeling, Regression Analysis, Statistical Analysis, SciPy, Pandas, Numpy, NLP

Soft Skills: Communication, Collaboration, Leadership, and Problem-Solving

Education

University of Sheffield, MSC in Data Science, Grade: 2:1

Sept 2023 – Sept 2024

- Data Analysis, Data Visualization, Data Mining, Database Design, Data Society, Business Intelligence.

Experience

Big Data Engineer, IQVIA – Bangalore

June 2022 – Aug 2023

- Collaborated with a team of 10+ members, optimizing ETL Spark using Scala and Airflow pipelines, reducing run time by 30%, and decreasing runtime from 10 to 7 hours.
- Enhanced efficiency in team workflow and automated redundant tasks using bash and Python, which increased operational boost by 10%.
- Automated airflow pipelines for cluster cleanup, resolving space deadlocks and receiving high appreciation.
- Worked on debugging and optimizing weekly crucial patient data pipelines to ingest terabytes of data using Airflow, Spark and Elasticsearch.

Machine Learning Engineer, CDAC – Pune

October 2021 – June 2022

- Facilitated the kanthasth project, involving the development of an AI chatbot using JS/NLP for Kantasth project, which reduced manual queries from users by 35%.
- Joined a 20+ team to create an NLP and text analytics dashboard for research purposes, reducing analytics time by 50%.
- Created Language Translation Models across Indic languages with an accuracy of 85% using transformers for languages such as Hindi and Marathi.
- Engineered and implemented frontend user interfaces for text analysis using Django and NLP, enhancing user accessibility and leading to a 25% increase in research analysis.

Academic Projects

Smart Cassava Leaf Disease Detection using CNN and OpenCV (Academic Research Project)

- Experimented with CNN architectures to create and test models with precision, AlexNet (68%), ResNet (55%) and ImageNet (78%).
- Trained models in iterations using MLOps practices, DVC (Data Version Control), CI/CD, Keras, and Tensorflow for GPU acceleration, reducing training time from 2-3 hours to 40-50 minutes on large datasets.

AI-Driven Stock Price Forecasting with LSTM and MLOps (Independent Project)

- Predicted stock prices with 92% accuracy, achieving RMSE of 5 by analysing 5 years of historical data.
- Reduced model training time by 40% using MLflow and DVC for versioning and tracking, enabling efficient deployment on AWS with 10x faster inference.

High-Precision Breast Cancer Detection using Ensemble ML Algorithms (Independent Project)

- Achieved 97% accuracy in breast cancer detection by applying Random Forest, Decision Trees (95%), and Logistic Regression (84%) to 570 patient records and 31 clinical features.
- Improved prediction time by 30%, outperforming baseline models in a Kaggle competition with a top 5% leaderboard ranking.