

# Kamesh Kotwani | Data Scientist

Sheffield, UK | Email: kameshkotwani@gmail.com | Phone: +447932851567 | LinkedIn: in/kameshkotwani

## PROFESSIONAL STATEMENT

---

Why hire two when one can do it all? As an Azure-certified Data Scientist and Data Engineer, I build scalable AI solutions, deploy production-ready ML models, and design robust data pipelines. I bridge the gap between data science and engineering, delivering end-to-end solutions to transform data into business growth.

## EDUCATION

---

MSC in Data Science | University of Sheffield | Grade: 2.1

September 2023 — September 2024

## TECHNICAL SKILLS

---

**Programming:** Python, Scala, R, SQL, Bash

**Cloud & DevOps:** AWS, Azure ( ML, Fabric, Synapse, Databricks), Docker, Kubernetes, CI/CD, Airflow

**Machine Learning Tools:** PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, MLflow, DVC, FastAPI, GenAI, LLMs, Statistical Analysis, Hypothesis Testing, Predictive Modelling, Visual Studio Code, Git, GitHub

**Data Visualisation:** Power BI, Tableau, KNIME, IBM SPSS

**Big Data Tools:** Spark, PySpark, Hadoop, MongoDB

## EXPERIENCE

---

TEACHING ASSISTANT | University of Sheffield

February 2025 — Present

- Coordinated in teaching **Big Data Analytics**, guiding students on **AWS, Databricks, PySpark, and Python**.
- Conducted lab sessions and tutorials on real-world big data solutions and improved understanding by **30%**.
- Provided support for students' practical sessions, including coding assignments in **Databricks** and **AWS**.

BIG DATA ENGINEER | IQVIA

June 2022 — August 2023

- Ensured data security and compliance with organisational data policies while handling large datasets.
- Built **Airflow pipelines** with Spark and Elasticsearch, improving platform reliability by **20%**.
- Automated **Spark SQL** tasks using Bash, increasing productivity by **10%**.
- Developed a Hadoop cleanup pipeline, reducing storage disruptions by **45%**.
- Optimised Spark runtimes by **30%** in collaboration with Spark and Hadoop teams.
- Enhanced CI/CD pipelines using Kubernetes, Docker, and GitLab, ensuring rapid deployments.

MACHINE LEARNING ENGINEER | CDAC

October 2021 — June 2022

- Led the **Kanthasth project**, developing an AI chatbot using JavaScript and NLP, reducing user queries by **35%**.
- Built an NLP dashboard with a team of 10+, cutting analytics time by **50%**.
- Developed translation models for Hindi and Marathi with **90% accuracy** using LLMs.
- Implemented **Topic Modelling** via Django and NLP, enhancing the efficiency of research by **25%**.

## CERTIFICATIONS

---

Microsoft Certified Azure Data Scientist Associate (DP-100)

Verify Certification

## MACHINE LEARNING PROJECTS

---

- **Sentiment Analysis with (FastAPI, Sklearn, DVC & MLflow) on AWS** Sentiment Analysis — Source Code (Logistic Regression, Count Vectorizer, GitHub pipelines, AWS(S3, EC2, FastAPI)
- **Comprehensive WhatsApp Chat Analytics** Chat Analysis — Live Demo
- **Smart Recommender Systems for Books and Movies** Recommender — Live Demo
- **High-Performance Stock Price Predictor, Time Series** (Achieved **90% accuracy**, **RSME 3.1**, using LSTM, GRU, 5 years stock data)
- **AI-Driven Breast Cancer Classification** (Achieved **97% accuracy** using XGBoost, Random Forest, Ensemble Learning)
- **Rapid Cassava Leaf Disease Diagnosis - (Research Dissertation)** (**93% accuracy** using Transfer Learning and CNNs, MLflow, GPU, DVC pipelines)