

SUJIT VARMA

+44 7741011522 | sujitvarma70@gmail.com | [LinkedIn](#) | [GitHub@Sujeet](#)

SUMMARY

MSc Data Science graduate from Queen Mary University of London with expertise in Python, R, SQL, statistical modeling, and machine learning. Hands-on experience in predictive modeling, big data analytics (AWS S3), deep learning, and AI agent development, with achievements like ranking at top 1% in CIFAR-10 Image classification (Deep Learning).

SKILLS

Core Programming: Python (Advanced), SQL, R, C#

Machine Learning & AI: PyTorch, TensorFlow, Scikit-learn, Deep Reinforcement Learning (PPO, SAC)

Specialized Domains: Computer Vision (CNNs, Attention Mechanisms), Game AI Development, Explainable AI (SHAP, LIME)

Data Science Stack: Pandas, NumPy, Tableau, Matplotlib, Jupyter, Predictive & Statistical Modeling, A/B Testing, Excel

Development & Deployment: Git, Streamlit, Unity ML-Agents, GPU Computing, AWS S3, Apache Spark

EXPERIENCE

Data Science Intern

Sep 2023 – Sep 2024

First Dot UK Ltd

Chester, UK

- Designed SQL database schema and wrote complex queries to extract and transform 10K+ customer records, integrating multiple data sources for comprehensive sales analytics
- Developed predictive forecasting models using both Python (scikit-learn) and R statistical packages, achieving 85% accuracy and enabling £30K+ revenue optimization
- Implemented advanced statistical analysis in R for customer segmentation and trend identification, combined with automated Python/Pandas data processing pipelines

PROJECTS

Computer Vision: CIFAR-10 Classification (Top 1% Ranking)

Python, PyTorch, GPU Computing

Queen Mary University of London

- Delivered state-of-the-art image classification system (93.99% accuracy) using enhanced CNN architecture, exceeding industry benchmarks by 8%
- Implemented advanced training strategies including data augmentation and OneCycleLR scheduling, reaching 99.77% training accuracy

Big Data Analytics Pipeline

PySpark, AWS S3, GraphFrames

Queen Mary University of London

- Built scalable data processing pipelines using PySpark and SQL to analyze 14M+ Twitter records, movie ratings, and taxi trip datasets stored in distributed databases
- Implemented complex SQL queries with JOINS and window functions for real-time streaming analytics, combined with advanced graph algorithms (PageRank, shortest paths) on large-scale transportation networks

Sales Intelligence & Forecasting System

Python, scikit-learn, Tableau

First Dot UK

- Engineered end-to-end sales analytics platform with automated data processing, predictive forecasting (85% accuracy), and interactive dashboards
- Generated £30K+ revenue optimization while reducing manual work by 60+ hours/month and improving operational efficiency by 15% through intelligent automation

Explainable AI for Predictive Analytics

Python, XAI, Feature engineering, scikit-learn

Freelance ML Consultant

- Built and deployed explainable ML web application using Streamlit, SHAP, and LIME, providing real-time model interpretability for business stakeholders with 83%+ prediction accuracy
- Validated model explanations through user testing with decision-makers, integrated feedback to enhance transparency and achieved 25% faster business decision-making

Game AI: Intelligent NPC Development

Python, C#, Unity ML-Agents

Personal Project

- Implemented deep reinforcement learning algorithms (PPO, SAC) to create human-like AI agents for video game environments
- Reduced model training time from 120 hours to 6 hours through reward engineering and model architecture improvements, validated performance through A/B testing with user groups and incorporated feedback to optimize model efficiency

EDUCATION

Queen Mary University of London

Sep 2024 – Sep 2025

MSc, Big Data Science

London, UK