

Neelansh Balhara

Lead Data Scientist | AI Engineer

Principal-level AI Data Scientist with 7+ years of experience building real-world ML and GenAI systems across healthcare-adjacent domains. Hands-on expertise in LLMs, NLP, and real-time inference pipelines, with a strong grounding in privacy-first design and regulatory compliance (GDPR, DPA). Built AI-powered GP Assistant and medical text automation tools. Passionate about accelerating clinical and digital health systems through responsible AI.



neelanshbalhara@gmail.com



+91 8178581329



New Delhi | London



project-nworks.vercel.app/



linkedin.com/in/neelanshbalhara-datascientist



github.com/neelansh22



SKILLS

Languages & Frameworks

Python · SQL · R · Groovy · ReactJS · PySpark · Tensorflow · PyTorch · NLP · TF-IDF · Signal Analysis

Data Technologies

ETL · Data Warehousing · Big Data Analytics · NER · Data Virtualisation · PowerBI

Deployment Platforms & Methodologies

Snowflake · Azure Web Apps · Docker · CI/CD · GitHub Actions · Azure DevOps · MLOps

AI Platforms

HuggingFace · MISTRAL Studio · LLM Fine-tuning · Azure Vision AI · Azure ML Studio · Azure AI Studio



EXPERIENCE

Lead Data Specialist | AI Engineer Valcon

01/2023 - 03/2025

London, United Kingdom

Data, Tech, and Consulting firm with 1,500+ employees across Europe, the UK, and North America, generating \$359M+ in annual revenue.

Achievements/Tasks

- Led the cross-functional requirements sessions, development and deployment of ML-driven personalised pricing engines for multiple enterprise clients, delivering optimal price corridors and increasing gross margins within 24 months of implementation.
- Consistently valued and scored three potential ML platforms (Azure ML Studio, AWS SageMaker, Databricks) against business benefit criteria, leading to a unified tooling standard across data science UK for predictive analytics projects.
- Delivered a customer segmentation model (rule-based + ML driven) in collaboration with marketing, operations and routing teams, for a UK-based energy giant enabling a 20% uplift in loyalty conversions in their national commercial market.
- Managed and mentored a growing team of data scientists; led deployment of a dynamic pricing engine for a Nordic aeronautics giant, enabling a positive uplift of 10% more engine part sales.
- Developed and presented a comprehensive business case for a leading UK automotive client—championing the implementation of a cancellation-prediction engine that drove a projected £1.4 million in gross savings p.a. by proactively identifying and mitigating high-risk bookings.

Data Scientist Valcon

01/2021 - 01/2023

London, United Kingdom

Achievements/Tasks

- Centralised 10+ PowerBI multipage KPIs reports by migrating complex DAX logic into a unified SQL Server data layer, significantly cutting report maintenance effort and reducing duplication of logic by 80%.
- Used NLP (TF-IDF + sentiment analysis) on news and media feeds to showcase business value extraction from a corpus of large natural texts for consumer insight generation, creating an in-depth analytical report regarding document confidence. Easily transferable to healthcare documentation.
- Represented my organisation at various GenAI capability showcases with scalable prototypes (Deloitte Challenge AI and Home Office Vendor's AI capability showcase)



EXPERIENCE

Data Consultant (Year in Industry)

Viqtor Davis

09/2019 - 01/2021

Sittingbourne, Kent, United Kingdom

Data and Tech house, serving UK's Automotive and Government sectors since early 2000s

Achievements/Tasks

- Achieved an average 90% accuracy on four Azure-based inference APIs forecasting workload for 300 garages over a three-year dataset.
- Accelerated query performance by 20% when migrating a 900 GB AWS Redshift warehouse to Azure Synapse, streamlining end-to-end ETL workflows.
- Reduced data inconsistencies by 35% across data accuracy and completeness metrics by designing a data quality scoring framework and mitigation roadmap for Unilever's Ireland operations.



PROJECTS

Project GPAi - GP Assistant AI (12/2024 - 02/2025)

- Python + Mistral + Streamlit | [GitHub](#) | [Live Application](#)
- AI-powered GP Assistant that clusters symptoms via vector embeddings and generates condition predictions and confidence-based narratives using LLMs. Built with data minimisation, real-time inference, and privacy-first design principles. This project has quite a bit of clinical use potential.

Project PIIT - Personal Information Identification Tool (05/2024 - 06/2024)

- Python + spaCy + Regex | [GitHub](#)
- Offline multilingual PII detection/redaction engine to secure sensitive medical and personal text data, enabling compliant NLP workflows.

Project IDOC - Intelligent Script Documenter (08/2024 - 09/2024)

- Python + Azure AI + Streamlit | [GitHub](#)
- LLM-driven documentation generator for clinical-like scripts and legacy systems (.py, .sql, .xml) with no persistent data retention.

Project OTIS (03/2025 - 06/2025)

- Python + Azure Vision AI + Streamlit | [GitHub](#)
- Image-based PII redaction engine that detects and scrubs sensitive visual text fields, forming the first layer of a privacy-focused OCR suite.

Project SDoD - Synthetic Data on Demand (03/2025 - Present)

- Python + Google AI Ecosystem + Private LLMs | [Github](#) | [Live Application](#)
- A complete end to end synthetic data generation suite, domain agnostic and supports private LLMs paving the way for privacy first AI design.



EDUCATION

MSc. in Advanced Computer Science (Computational Intelligence)

University of Kent

08/2018 - 08/2020

Canterbury, United kingdom

Thesis

- [A Critical Analysis of Facebook-Cambridge Analytica scandal, Regulation of Data protection legislation \(UKs DPA 2018 & GDPR\) and the future of Data-Driven Political campaigns.](#)



CERTIFICATIONS

- Microsoft Azure Data Scientist Associate (In Progress)
- IBM Watsonx Foundations (Planned)



CORE COMPETENCIES

- Cross-Functional Collaboration
- Problem Discovery & Framing
- Mentorship & Uplifting Others
- Stakeholder Engagement
- Empathy with End Users
- Strategic Customer & Behavioural Analytics
- End to End User Experience
- Data Engineering & Cloud Architecture