

Aatish Thakerar

London, UK | +44 (0)330 001 0390 | aatish@theaiguy.tech | [Linkedin](#) | www.theaiguy.tech



PERSONAL SUMMARY

A client-facing Data Scientist knowledgeable in the domains of Machine Learning (ML) and Deep Learning (DL), specifically Data Mining, Computer Vision and Natural Language Processing (NLP). Primarily product-focused, having built interoperable technologies from Prototypes to Production-level deployments. Commercially-astute in using statistical analysis to discover client issues and targeted workshops to resolve them, across a number of product and service landscapes in disparate industries. Possesses excellent stakeholder management and workflow coordination skills in often pressurised environments.

TECHNOLOGIES & FRAMEWORKS

(c. 2015) **Python:** Pandas, NumPy, Scikit-learn, TensorFlow, Keras; **UNIX/Linux:** Bash/Shell scripting
(c. 2016) **Data Visualisation:** Tableau, Microsoft Power BI; **Container Deployment:** Docker, Kubernetes; **Git; Wireshark**
(c. 2017) **SQL:** MySQL, PostgreSQL; **NoSQL:** MongoDB, Apache Cassandra, Neo4j; **Java; C; Splunk**
(c. 2018) **Apache Big Data:** Spark, Hadoop; **Amazon Web Services (AWS):** Solution Architecture, Machine Learning

LEADERSHIP

- Performance Manager for several Consultants at Capgemini, Technical Delivery Lead on R&D-based client projects
- Leading workstream of R&D, overseeing multiple Data Scientists, Data Engineers and Software Engineers at WiiA
- Managed workflow/project slates worth cumulative total of ~£7 million from 2013–Present day
- Trained specialist teams on using workshops to turn raw objectives/requirements into technical specifications
- Mentored six graduates in 2018/19 and three graduates in 2019/20 across Accenture – all were promoted

CAREER HISTORY

Capgemini Invent, London – Senior Consultant, Analytics and AI

Oct 2020 – Present

Invent uses multi-disciplinary teams to accelerate ideas into prototypes and scalable real-world solutions

- Technical Delivery Lead for Innovation Lab at UK Welfare Department, focused on 6-week to 6-month technology explorations: NLP Summarisation, Personal Data Stores (Digital Identity), Synthetic Data, Explainable AI / Ethical AI
- Co-led a research collaboration with the Alan Turing Institute to explore the privacy characteristics of Synthetic Data
- AI-Powered Innovation: Thought/team leader on the go-to-market strategy of scaling AI solutions up from PoC
- Workstream lead for ‘AI Playbook’: valuing the proprietary AI technologies the Capgemini AI team offer to clients

WiiA, London – Co-founder & Chief Technology Officer

Mar 2020 – Present

WiiA is an advanced Engineering Sector start-up aiming to address inefficiencies within the tender and proposals market

- Steering development of an AWS-based, NLP-powered product, guided by an 18-month PoC-to-Production roadmap
 - Using funding to continuously iterate WiiA’s algorithm and product; to meet our goal of automating 80% of proposal content through AI, we baseline performance against multi-purpose models BERT & GPT-3
- Working closely with 3 Multinational Engineering clients to develop a precise & real-world relevant solution
- Reporting to Board of Advisors including current, senior figures from industry (Microsoft, Atkins, Accenture)

The AI Guy Ltd, London – Founding Director

Mar 2020 – Oct 2020

Contracting company founded to bring a more logical approach to AI for clients

Lead Data Scientist at Engineering Consultancy providing digital integration services

Mar 2020 – Present

- Utilising semantic image segmentation, a Computer Vision technique of supervised learning, to correctly label ~70% of street-level image data from sites across developing countries, and at < 5% the time taken for manual methods
- Applying object detection and classification methods to geospatial aerial image data, marking smaller road-like surfaces and predicting the feasibility of running local transport infrastructure routes on them
 - This AI method of surveying has so far replaced ~20% of manual visits to these sites across 3 countries

Accenture, London – AI Specialist, UK Cyber Security

Oct 2017 – Mar 2020

Held positions as Cyber Security Analyst, Cyber Security Consultant, and Analytics/AI Engineer

Lead client-side Data Scientist for Zoran, a predictive modelling tool to map an organisation’s Identity Access Management (IAM) landscape, and forecast the required levels of user access in a system/application

- Wrote a custom Association Rules (AR)-style mining algorithm, based on a novel Regression-type analysis of users' attributes and levels of access; this approach's feedback allowed us to automate feature selection in model training
 - Using this unsupervised learning heuristic, we discovered user-access relationships (at minimum 80% statistical confidence) 6-8 times faster and 30% more accurately than a dedicated human IAM team
- Iterated from R&D to MVP using Agile-style Sprints with regular Scrums, supplemented by JIRA & Confluence to manage workflows, and a CI/CD methodology to improve the tool via its modular microservices architecture
- Built data pipelines to write AR-generated rules to Cassandra as Parquet files, transform them to CSVs using Pandas, and pull them to the front end to populate both Zoran's UI and our custom Tableau views and Power BI visuals
- Led teams of Data Scientists & Data Engineers to implement the tool across client environments – turning around a PoC in 5 weeks to a competitor's 10 months, and refactoring complex code to reduce model training time by 90%
- Across all client deployments, Zoran automated the processing of 40-45% of user access requests, scaling successfully to systems with 500,000 users and 2 million user access roles
- Zoran was acquired by ForgeRock, the leading IAM provider, in December 2019, and assimilated into their new ML-backed offering called 'ForgeRock Autonomous Identity'

Specialist Integrator of 3rd party DL platform for 2 world-renowned Law Enforcement agencies tackling gang crime online

- Wrote a Java-based tool to assess the statistical strength of results from disparate DL engines, pull together their best segments, and produce a single coherent output for OCR, transcription, object detection and facial recognition
 - Results were, at a bare minimum, 6% more accurate in aggregate form than any single engine in isolation
- Created new metrics to measure DL engine accuracy. For example, 'Context & Key Word Rate' was designed to measure Automatic Speech Recognition against specific law enforcement objectives in verbal interrogation scenarios; it registered an average accuracy rate of 75%, compared to generic metric 'Word Error Rate' at 60%
 - Built Python-based scripts to automate the accuracy scoring for DL engines against the new metrics
- Law enforcement used the integrated, optimised platform to automate the search for, flagging of, and take down of incendiary gang videos on social media within half a day, compared to a prior 3-week-long manual method

Led experimental projects to draw new analytics techniques from emerging technologies, for example, the use of MIT's probabilistic multivariate database BayesDB, to further differentiate Accenture's data-driven services in industry

Built strategic UK Security Practice frameworks 'AI in Security' and 'Securing AI', managing a team of 3 Consultants to both benchmark clients' AI readiness and aid them in creating automation roadmaps across a total of 22 projects

GfK, London – Technology and Business Analyst

Oct 2014 – Apr 2016

- Developed a Robotic Process Automation (RPA) powered methodology of acquiring high-level, actionable insights from raw data reports, reducing the amount of time scouring raw data by 80%
 - Stored key insights in a dedicated *Amazon RDS for PostgreSQL* database, and created a selection of pre-defined, commonly used SQL queries for non-technical users to be able to extract the most recent data
- Conducted market segmentation across complex datasets, using custom macros, SPSS and modules for analysis in Python and R – identifying a further £800k of commercial opportunities for GfK in the consumer technology sector

Campden Wealth, London – Research Executive

Aug 2013 – Oct 2014

- Executed the day-to-day research for 6 Financial Services reports, with a total contractual value of £1.8 million
- Conducted interviews with 70+ stakeholders from service providers, family offices, and subject matter experts
- Co-authored three quantitative research reports, including Campden's largest global project on wealth management
- Built and deployed an in-house Analytics platform across own-brand sites, increasing ROI by 23% and 32% for two regional reports, and 144% for the flagship global report

EDUCATION AND CERTIFICATIONS

Queen Mary University of London – MSc Computer Science (Distinction)

Sept 2016 – Sept 2017

Thesis: 'Neural Networks for Reinforcement Learning in a Multi-Agent Environment'

Description: Applied Reinforcement Q-Learning via a neural network 'hive mind' to boost the collective efficiency of collaborative agents in an ant colony simulation (*awarded Distinction*)

The Open University, London – Essential Mathematics 1 & Essential Mathematics 2

Sept 2015 – Jun 2016

Fundamental topics underpinning university-level study in mathematical subjects

University of Nottingham – BA English (2:1)

Sept 2010 – Jun 2013

Thesis: 'Analysing modes of post-feminism in the gothic narratives of Angela Carter'

Certifications:

- Amazon Web Services (AWS): Certified Solutions Architect - Associate Nov 2019 – Nov 2022
- Market Research Society: Advanced Higher Certificate In Market and Social Research Practice Jun 2015 – Present