Adam Fletcher, PhD

dradamfletcher@gmail.com +44 (0) 7707722774 adamfletcherUK adamfletcherUK

I am an enthusiastic data scientist with experience in both academic and business settings. My academic career utilised genetic data, developing tests to detect Down syndrome in expectant mothers and characterised cancer patients' response to chemotherapy. In business, I have worked for multiple clients helping them solve problems with data and improving their data-driven business decisions.

KEY SKILLS AND EXPERIENCE

Data Science:

- Building end-to-end data pipelines handling: ETL, feature engineering, model training and API deployment.
- Utilizing clustering to solve problems such as personalizing web services to predicting chemotherapy response
- Topic modelling using BERT to cluster conversation transcripts, surfacing common questions and pain points.
- Using regression models to accomplish tasks such as: predicting quality changes over time, calculating machine sales prices and normalising pricing data as part of a price index.
- Created a question answering webapp utilising RoBERTa, where users could ask a question on a store of unstructured documents. Outputting an answer, contextual information, and a confidence score.
- Generating accurate synthetic data using GANs (Keras) for testing data migration services requiring realistic data.
- Promoting ML explainability using SHAP and LIME to enable end-users to determine why a prediction was made.
- Timeseries forecasting using 1D convolutional networks and LOESS regression to predict future trends.
- Created a multi-phase financial model for project costings, capable of predicting monthly pricings, initial debt, profit, and growth rate and how this is affected by payment terms for a project.

Data Engineering:

- Migrating Jupyter notebooks to production CI/CD data pipelines, complete with ETL, feature engineering, model training, deployment, and performance reporting.
- Deploying production ML APIs with multiple endpoints, serving model predictions for different purposes.
- Building simple data lakes on AWS and using AWS Glue crawlers to query multiple files and databases.
- Utilising frameworks like Dask and cuDF to parallelize and speed up computation time on large datasets.
- Experience using: Python, Jupyter, R, SQL, Shell, Tensorflow, Git, docker, Dask, Flask, Travis, AWS Services, Heroku.

Communication:

- Possess outstanding written and communication skills; disseminating technical information to board members, ensuring conclusions are comprehensible to technical and non-technical personnel.
- Extensive remote working experience as part of multinational teams or with no geographical location

RELEVANT PROFESSIONAL EXPERIENCE

Consultant Data Scientist Equal Experts **2020 - Present**

Predicting the resale price of used machinery

- The client wanted to determine factors that influenced the resale price of machines to better optimise sales prices.
- Created a non-linear regression model to predict the price of an asset and how that price changes over time, which was deploying via a Flask API.
- Conducted validation activities to build user-confidence in predictions as the industry was typically not data-driven.
- Built a pricing index, modelling market trends over time, using regression to normalise sales to reference values.
- The business is now able to sell assets faster and at a more appropriate price. They are also able to predict depreciation price on new assets and factor in price of selling assets in project costings.

Topic Modelling Call Centre Transcripts to Determine Common Conversations

- The client wanted to surface common questions being asked in their call centre, to better serve customers.
- Utilised the BERT tokenizer to embed the call transcripts then performed dimension reduction and clustering to identify common themes and topics in conversations between call centre staff and customers.
- The business is now able to look at common questions, common topics and questions around those specific topics.
- The client can now surface this information to customers faster.

Demonstrating AI Potential in a Large Government Organisation:

- Timeseries forecasting using 1D convolutional networks to analyse multivariate temporal data.
- Using GANs and statistical modelling to create realistic synthetic data for use with testing data migrations.
- Text classification using SpaCy and random forests for predictions and LIME to explain the predictions.
- Built a graph database in GPU memory (cuGraph) of company's house data to show a network of companies and their major controlling parties for querying.

Data Analyst Equal Experts **2019 - 2020**

Creating needs-based personas for a job-listing service:

- The client wanted to understand how users interacted with a job-listing website, aiming to personalise user experiences based upon their needs.
- As the survey dataset was incomplete, we used matrix-factorisation to predict user responses for unanswered questions, followed by clustering to group users with similar needs.
- Clusters were used to make personas with user needs and the business was better able to respond to customers.

Modelling the cost of debt for delayed payment terms in commercial projects:

- The client wanted to understand how the debt resulting from delayed payments terms affected the final profit of a project and company liquidity.
- Using incoming and outgoing costs for each project, I created a multi-phase model for running projects, including the amount of debt accrued and how increased payment terms affected debt the cost of credit.
- With this information, the client restructured their pricing system and was better able to forecast their liquidity at key times of the year.

Reducing the cost of variation in providing a service:

- The client wanted to restructure their national pricing model. But each franchise provided the same service with minor changes, resulting in differential cost to the business.
- Working with a new workflow tracking system, we were able to attribute normal operating criteria and areas which were operating outside of them.
- In a detailed report outlining standard operating procedures I also determined the cost-saving resulting from implementing the new procedures, which exceeded £2million.

Postdoctoral Researcher Cancer Research UK **2017 - 2019**

Identifying markers leading to chemotherapy resistance in small cell lung cancer:

- The aim was to develop a blood test, assessing and predicting patients' response to chemotherapy.
- Analysing multiple complex datasets to identify features driving drug-resistance, accounting for the large patient-specific genetic variations seen in biological datasets.
- Processed genetic data and performed clustering to group pre-treated patients as drug-sensitive or resistant.
- Utilising R to design data analysis pipelines for novel laboratory methods I have developed.
- The outcomes included determining drivers for chemotherapy resistance which could prolong the lives of patients.

Molecular Diagnostic Scientist Premaitha Health Ltd. **2014 – 2016**

Developing DNA prenatal screening tests for Down Syndrome:

- Developed blood tests that predict the likelihood of Down Syndrome in the foetuses of expectant mothers.
- Determined quality statistics which affected the false-positive and negative rates of the test and altered acceptance criteria for the test.

Using timeseries forecasting to determine manufacturing quality statistics:

- Developed a post-market surveillance system to establish and monitor quality metrics for the commercial products.
- Analysed metadata from internal and customer data to establish suitable quality thresholds for routine data monitoring and whether these were drifting over time.
- The outcome was the ability to forecast product quality drops and alert the support team for intervention.

EDUCATION

PhD in Molecular and Structural Biology University of Manchester **2011 – 2016**

- Aiming to improve a chemical reaction by understanding the molecular interactions occurring and enhancing functionality.
- Developed and built 3D structural models of novel proteins and interrogated their chemical binding through experimentation and simulations.
- I was able to sufficiently alter the binding properties enough to change the target chemicals in the reaction

BSc. (Hons) Biology University of Manchester **2007 – 2011**

Result: 2:1

PERSONAL

I love competing in sports, playing Roller Derby for over a decade, representing England and competing in world championships. I frequently tinker with electronics with a raspberry pi and have built an arcade cabinet.

REFERENCES

Available on request