Matthieu Marshall

Tel: 07789843468, Email: matthieu.marshall@gmail.com

Data Engineer with key skills in data wrangling using Python, SQL and Pyspark. Currently working for a market research company using Azure Databricks and Data Factory among other technologies for ETL processes. Previous experience at a technology consultancy working with Python, Scala, SQL and elements of the Hadoop stack for the ingestion, processing and surfacing of data.

Experience

Data Engineer at Redslim *(July 2020 – Present)*

- Used Azure Databricks to clean and process sales data into a standard format.
- Migrated ETL steps from SQL to Pyspark.
- Coordinated ETL processes using Azure Data Factory.
- Worked with Azure DevOps to manage work items and administer CI pipelines.

Data Science Engineer at Roke Manor Research Ltd *(November 2017 – July 2020)*

I started at Roke on the graduate development program, which I finished in November 2019 when I was promoted to a Data Science Engineer. During this time, I was involved in the following:

Project supporting development of a Big Data Pipeline for use in the defence domain.
(February 2019 - Present)

> **Technical domains**: Scala, Python, Git, Shell, SQL, Apache Hbase, NiFi, Kafka and Hive, HDFS, Docker, Groovy & Jenkins

- Developed a data pipeline to ingest data which is stored and then presented to the user using a GraphQL API.
- Worked on customer site in a rainbow team alongside other contractors.
- Shadowed a data architect role to produce documentation for accreditation.
- Created a pipeline for ingest of video over UDP stream using Gstreamer in Python.
- Experienced working in an agile scrum environment with continuous integration and behaviour driven development, managed via Jira.
- Communicated with users and external stakeholders to explain the platform and support their integration to it.
- Worked in a secured Linux environment, using shell commands day to day.

Project providing test support to a database development project in the law enforcement domain.
(March 2018 – February 2019)

> **Technical domains**: Python, PostgreSQL, Git.

- Delivered test data for migration of and creation of a new law enforcement database.
- Risk assessed migration from Microsoft SQLServer to PostgreSQL.
- Used Python, SQL and custom-made software to create and maintain a range of complex test data in a PostgreSQL database.
- Wrote custom SQL queries to validate test data.

Project conducting a capability assessment of open source methods for face recognition.
(November 2017 – January 2018)

- Benchmarked open source face matching tools in Python using packages such as sci-kit learn and opencv.

- Investigated and carried out web-scraping for labelled face images.
- Trained the best tool found to successfully identify individuals from a photo.

Project assisting a cyber vulnerability assessment in the automotive sector.
(January 2018 – March 2018)

Technical domains: Python, bus networks, modern automotive concepts.
- Investigated the security of the internal network of a modern car.
- Used Python to collect and visualise bus network data to find potential vulnerabilities in its' design.

Other work activity:

- Contributed to a business development task to produce a Python tool to augment audio data for Machine Learning purposes.
- Assisted with the organisation and running of workshops with defence stakeholders.
- Successfully supported Roke in bidding for a Defence and Security Accelerator project.
- Planned and delivered activities for STEM promotion events and school visits.

Education

The University of Southampton *(September 2014 – June 2017)*
   BSc Mathematics with Statistics, First Class Honours

- Completed a group project to analyse airline fares using multivariate regression in R to verify a claim about their characteristics.
- Produced graphs of financial, statistical and population data to complete coursework.
- Learnt the theory behind and put into practice regression, factorial and survival analysis techniques.
- Learnt about the properties and applications of a range of statistical distributions.
- Volunteered at CodePlus, and took on the role of President of the University athletics club.

Online Courses completed
Advanced SQL for Machine Learning from LinkedIn Learning.
Advanced Python from LinkedIn Learning.
Using Power BI with Excel from LinkedIn Learning.
Understanding Machine Learning with Python from Pluralsight.

Cherwell School and Sixth Form, Oxford

Studied for A Levels in Mathematics(A), Further Mathematics(B), Geography(A) and French(B)

Awards & References

Award for work on Programme NELSON while at Roke Manor Research Ltd *(April 2020)*

"Matthieu ... ended up working most of the weekend to resolve issues. He did this off his own back which highlights the level of professionalism and dedication he always applies, and why he is one of the strongest people we have on the programme.... I can't thank him enough for his efforts."

This is an outstanding effort, particularly during the current times. Matthieu is a very well-respected member of the team, and recently he has been taking on more responsibility and using his own initiative to solve problems as they arise.

Languages

English: Fluent   French: Fluent   Spanish: Basic