**Shuqing Ke**

M: 07957200696 /A: 106 Southover, BR1 4RY /E: shuqingke1404@gmail.com

An innovative, result-oriented and highly motivated data scientist with a strong Python, PySpark & SQL quantitative analytical background and proven track record of deriving actionable data-driven insights.

**Employment History**

**(11/2023 - 10/2024) Contractor, BP**

* **ETL Development and Data Integration**: Developed robust ETL processes integrating diverse data sources (PROS and InsightHub) using Python, SQL, and AWS.
* **Join Logic Validation:** Established secure connections and performed data cleaning and feature engineering to enhance data integrity. This involved addressing complexities such as variations in attribute names and reconciling multiple delivery points associated with a single shipping point. Implemented validation checks to prevent duplicate price entries in PROS, mitigating the risk of erroneous joins with transaction records
* **Visualisation Development:** Created a Python MVP dashboard and contributed to the PowerBI visualisation, facilitating improved pricing strategy decision

**(12/2021-09/2023) CGI Inc., Data Scientist**

* Conducted statistical analyses (ANOVA, Kruskal–Wallis) and utilised time series models (ARIMA, Prophet) to forecast sales

**(04/2022 – 04/2023) Shell, Data Scientist Contractor**

* Implemented production codes for the Shell Offer Decision Engine in PySpark
* Implemented new feature such as B2B/B2C flag as per business requirements in PyCharm and conducted simulation testing by creating simulated database and pipelines in Azure Data Factory
* **Skills learnt: PySpark, PyCharm, Databricks, Azure Data Factory, Git, simulation testing, Azure DevOps**

**(03/21-12/21) Data Scientist, Utility Warehouse, London**

**(03/20 - 03/21) Data Analyst**

* Involved in all the stages of Data Science cycle from understanding the business requirements to testing and optimising the solution within energy customer sector
* User Segmentation:

1. Use unsupervised learning(clustering) techniques to segment user base according to different feature sets
2. Use dimensionality reduction (Umap /TSNE) to avoid multicollinearity and help demonstrate clustering visualisation
3. Example project 1: clustered customers’ quantitative survey results using K-means and Umap to help build customers’ persona and new business value proposition
4. Example project 2: used cosine similarity to pick top 10% customers that are most likely to be receptive to upselling broadband products, and combined with geographic data using Folium to identify the customers who are within 10km distance from the partner base

* Predictive Modelling:

1. Use supervised learning (regression/classification) techniques in a variety of prediction tasks
2. Use Partial Dependency Plot, SHAP, survival curves and Plotly to better visualise features’ contributions to model’s predictions
3. Example project 3: developed an energy churn model with architecture based on random forests, and performed feature extraction using unstructured web data, statistical features and aggregations
4. Example project 4: developed a customer acquisition model (80% recall and 50% precision on validation set) and used statistical approach to rebalance training set, and visualised acquisition funnels and onboarding journeys to help support marketing and acquisition teams

* Pipeline Productionisation:

1. Involved in the deployment of ML pipelines using BQ Dataform to create and transform raw data into tables & views
2. Example project: developed a K-means BQ ML model, built data pipelines in Dataform and output the metrics results in Looker dashboard

* Production Model Data Drift:

1. Designed and implemented data input and model output drift detection using statistical tests, worked closely with data engineers to productionaise the results in Big Query
2. Based on the drift results, built a Looker dashboard to visualize the outcome to stakeholders
3. Example project: developed data input and model output drift for K-means model using statistical tests such as T test, the two sample Kolmogorov-Smirnov, Chi-squared test,Cochran’s Q test and Wilcoxon test to detect data drifts and trigger alerts accordingly

* **Skills learnt: Python, Big Query, Looker, Google Cloud Platform, Dataform**

**(07/19-03/20) Junior Data Scientist, Beamly Ltd (part of Coty Inc.), London**

* Prepared data and conducted quantitative analysis to build predictive models identifying consumer behavior patterns.

**(04/19-07/19) Consumer Journey Analyst, Coty Inc., London**

* Analyzed eCommerce growth priorities and created actionable reports for sales teams

**(09/16-04/19) Search (Digital Data & Insight) Analyst, Contrado Imaging Ltd, London**

* Worked directly with the director to undertake ad hoc insights analysis and deliver regular monthly KPI reports based on KPIs (revenue, consumer shopping behavior) via Elasticsearch & Power BI, presenting data informed actionable plans to stakeholders
* Continuously analyzed the website in terms of customer journeys, conversions via Google Analytics,, creating customized dashboards to improve conversion rates and ROI with increased customer acquisition, retention and sales of 20% YOY

**Education**

(2019) General Assembly, London

**Data Science**

(2018) Training Square, London

**Microsoft Certified Solutions Associate (MCSA) SQL 2016 Business Intelligence (BI) Development certification**

(2015-2016) Quartic Training, London

**Chartered Financial Analyst:** passed level 1, CFA program (5/12/2015)

(2013-2014) University of Glasgow, Glasgow

**MSc: Management with Enterprise and Business Growth** (Grade: Pass)

(2008-2012) Shenyang Jianzhu University, China

**BA: Human Resource Management** (Grade: 2:1)

(2005-2008) Fuzhou No.1 Middle School, China

Modules studied: Maths (A), English (A)

**Extracurricular Activities**

**(06/19-03/2020) Welcomer Volunteer, The Design Museum**

**Interests and Other Competencies**

**IT Skills:** Microsoft Office, Python, Pyspark, SQL

**Languages:** Mandarin (Native), English (fluent)

**Interests:** reading, hip-hop, house & locking dance, art galleries and volunteering

**References**

Available upon request