

Han ZHOU

han.zhou@hsbc.com | (+44) 07536403853 | Homepage: <https://hanzhou925.github.io/>

Education

MSc. in Language Science (Neuroscience), University College London Sept.2022
B.A. in German, Shanghai Jiao Tong University (Grade: 88/100) | Minor: AI Sept.2018-Jun. 2022
Courses: Natural Language Processing (90); Probability and Statistics (88); Introduction to Calculus (97); User-Centred Data Visualization; German Linguistics (94); Frontiers in Linguistics (98); AI Programming Framework; Introduction to AI **Awards:** SJTU Scholarships (Top 20%); Third Prize of National English Competition for College Students (Top 5%); IBM Certificate of Data Science

Work Experience

CMB Graduate Program, HSBC Jul.2023- Present

Rotation 2: Risk Data Analyst, Wholesale Data, Analytics and CRM

- Leading the project of growing customer safely using BigQuery to analyse and visualize customer data, thus identifying the industries driving the greatest growth for HSBC and the financial/non-financial risks alongside, and providing RMs with advices to mitigate risks for customers;
- Supporting the internal data analytical teams to migrate datasets and codes from SASS to GCP (Google Cloud Platform), thus updating the technical level and platforms within the bank.

Rotation 1: Relationship Manager, Business Banking

- Monitored and mitigated financial risks for SME clients by reviewing and reporting on their annual accounts and MI, as well as analysing their credit risks, CRRs and loan products to decide on waiving the quarterly MI;
- Aligned on the AD's portfolio management and migrated customers to BBD/BBPM by analysing their business accounts, credit risks and loan limits, etc.;
- Led the organization of all the new customers' portfolios in H2 2023 for the team by reaching out to customers and wider teams, thus designed and publicized the deal flyer for marketing on LinkedIn, receiving thousands of viewing;
- Collaborated with North and South Yorkshire RM teams and sustainable finance team to host upskilling sessions among local teams, thus enhanced the RM's knowledge about sustainable products and Sustainability Tracker.

NLP Product Operation Intern, AI-LAB, ByteDance Technology Co., LTD Feb. 2022-Jun. 2022

- Led the TikTok Cross-Language Assessment (CLA) project in German, cooperated with other language teams, and analysed data from over 1000 problematic captioned videos, thus determining factors affecting users' satisfaction with the live subtitle function on TikTok and delivering the outcome to R&D department
- Organised interviews with 5 groups of German speakers, researched their habits of using subtitles watching videos in foreign languages on TikTok, locating users' pain points and priorities to improve TikTok's subtitle function
- Analysed and visualized data from over 1500 corpus items from 5 dimensions, thus producing structured reports on the existing problems and providing feasible suggestions in communication with R&D to upgrade Feishu.

TTS Linguistic Data Analyst, Ximalaya Technology Co., LTD (top Internet company in AI voice with over 2000 employees) Jun. 2021-Sept. 2021

- Led the high-prioritized sentiment annotation project, analysed sentimental polarities and 8 categories of emotions in dialogues, thus establishing the sentiment dataset with over 24,000 data to develop Text-To- Speech (TTS);
- Manipulated over 70,000 polyphone Mandarin data, and added over 3000 missing data from BCC corpus, thus expanding the polyphone coverage of the dataset by 2% and improving the accuracy of machine learning through cooperation with R&D;
- Evaluated and fixed over 1000 bad cases of AI pronunciation in Mandarin based on prosody, phonetics and sentiment, regularised and cleaned the newly added Mandarin data with Python, enhancing the performance of speech synthesis

Projects

Data Science & Analytics Virtual Intern Program, BCG Jul. 2022-Aug. 2022

- Applied CRISP-DM Model to diagnose the source of SME customers for PowerCo., and put forward the hypothesis that churn is driven by customers' price sensitivity, locating the goal of modelling
- Performed exploratory data analysis based on over 15,000 data from clients, so that assisted the client PowerCo. in analysing how different factors like price sensitivity affected the churn rate of their customers
- Engineered the features after cleaning the datasets and built a Random Forest Classification model to predict which customer subscribed at PowerCo. is more likely to churn, reaching an accuracy of 90%

Image Restoration based on Linear-regression Analysis, SJTU

- Generated damaged image by adding noise masks to the original image, segmented the image into regions to apply Linear Regression Model in Sklearn and restore the image with a Cosine similarity of 0.998 with the original image

Feature Face Recognition (Eigenface) based on PCA, SJTU

- Processed face images from ORL database, applied covariance matrix to quantitatively describe "face difference" and create eigenface matrix, trained data with eigenface and average face and reached an accuracy of over 88% in prediction
- Transformed data into the eigenface space and inversely transformed it into the pixel space to rebuild human face

Other Information

Tech Skills: SQL; Python(sklearn/pandas/numpy); RStudio; GCP(BigQuery); Tableau; Looker Studio; Qilksense

Language Skills: Chinese (Native); English (Fluent); German (Fluent); Sanskrit (basic)

Soft Skills: Interpersonal Communication; Adaption to multicultural atmosphere; Teamworking; Learning capacity; Multitasking