

Chenghan Song

aliciasong95@outlook.com

EDUCATION

University of Amsterdam, Vrije University of Amsterdam

MSc in Computer Science, Big Data Engineering

Amsterdam

2020 - present

Soochow University

BEng in Computer Science

Suzhou

2013 - 2017

EXPERIENCE

Yipingfang

Software Engineer Intern

Nantong

March 2019-September 2019

- Utilized expertise in Scrapy framework and Python to crawl websites and extract data.
- Participated in the development of similar image search using MXNet and FAISS.
- Used Bootstrap and AngularJS framework for front-end styling and user interaction for testing.

Dataondemand

Business Intelligence Engineer

Shanghai

July 2017-November 2017

- Liaised and coordinated with the units of a Biotechnology Company to define data standards, determined business data specifications, and calculation logic.
- Designed and implemented a self-service analysis platform, mobile phone dashboard, and big-screen dashboard pages to present data visualization for different users.

PROJECTS

• Visualization of Covid-19 Flight Impact

- Pre-processed the 200 GB of compressed ADS-B messages from *OpenSky Network data* and *Covid-19 flight data* by decoding, extracting, and filtering position messages.
- Simulated flight patterns of the covid-19 data to obtain the aircraft position by matching the flight information.
- Created a visualization to display the Covid-19 flight impact and the density of the flight network quantitatively.

• Sentiment Analysis of 2020 US Election Tweets

- Pre-processed the data by pruning and cleaning the tweets ranging from 15.10.2020 to 04.11.2020.
- Constructed a bi-directional LSTM (Bi-LSTM) based deep learning model to take advantage of Word2vec word embedding for the sentiment analysis task.
- Investigated how does social media reflect the process and result of the election by visualizing the most concerned tweets topics, geography of political participation, and results of sentiment analysis.

• Tourist Destination Recommender

- Conducted visual analytics on the Flickr dataset and built a tourism destination recommender where images from the dataset could be shown to the user of possible travel destinations.
- Learned user preference by dividing the photos into groups and calculating the popularity score of each photo for each group.

• Expedia Hotel Recommendation System

- Processed the data by handling missing values and outliers of the original data. Conducted feature engineering and selection to help the model find latent patterns by utilizing the feature importance function provided by LGBMRanker.
- Applied LGBMRanker as the core model and employed LambdaMART to rank hotels by the likelihood of users clicking.