Jiale Dai

(646)920-1765
jiale.dai@nyu.edu
https://github.com/JialeDai
225 Cherry St, Apt 11B, New York, NY

EDUCATION

New York University	MS.Computer Science	GPA 4.0 /4.0	09/2020 - present
New York Institute of Technology	BS. Computer Science		09/2019 - 06/2020
Nanjing University of Posts and Telecommunications	BS. Computer Science		09/2016 - 06/2019

PROFILE

- Objective: Seeking for Software Engineer intern position.
- Language: Proficient: Java, Experienced: Python, C++, SQL.
- Technical Skills: Spring framework, Spring MVC, RESTful API, Maven, JUnit, SQLite/MySql/MongoDB/Redis, Git, Linux, MapReduce, HDFS, Elasticsearch, PyTorch, HTML, JavaScript, CSS, vim.

PROJECT EXPERIENCE

Sentiment Analysis of Coronavirus Base on Social Media Data in China 09/2020-11/2020

Website: *https://limitless-harbor-77448.herokuapp.com/*

- Designed a website to representing analysis result of people's attitude and the trend of coronavirus based on the posts from social media using python **Django** and combined the development of pandemic, the topic model to explain our prediction.
- Collected post information for the past four-month, which is roughly 600,000 unprocessed posts data by writing web crawler using **BeautifulSoap** and **Selenium**.
- Handled the big scale of data with noise using python and regular expression and store the cleaned data into **MongoDB** database.
- Fitted the processed data into different classifiers (SVM, Naive Bayes, Linear Regression) with different number of features and scale of data set, then chose the classifier with the best performance.
- Visualized the Covid-19 data by building dynamic heatMap which represents infection information by day using eCharts.
- Created the topic model and built the word cloud to show the top 50 hot topics on social media.
- Built data dashboard to explain the correlation between infection data and sentiment analysis result.

Text Generation Dialogue Bot 06/2020-09/2020

- Built an auto-response chat-bot website based on users' input text using **Spring Boot, Elasticsearch, PyTorch, Redis, and MongoDB**.
- Utilized Docker to set up the program development environment.
- Developed **RESTful** backend system to collect user input text and interact with **MongoDB** and machine learning model.
- Achieved complex data retrieval features using **Spring Data JPA and customizing JPA repository**, supporting data query by any data field, sorting order, return limit, and value range.
- Traversed the corpus for the existing answers based on TF/IDF using **Elasticsearch**.
- Improved data retrieval performance by using **Redis** as cache to store the most frequent conversations.
- Trained the Dialogue Model Based on the models in Microsoft(DialoGPT) and FaceBook(Wizard of Wikipedia)' s paper.
- Using **Socket** to get the predicted result from python program if the server cannot find a reasonable answer in cache and database.

Yathzee Program on Android Platform 03/2020-06/2020

- Developed a Yahtzee dice game on Android platform using **Spring** , which supports saving the unfinished game and loading/updating the stored game.
- Designed the GUI of the game using XML and implemented the internal logic of the program based on the game rule using Java.
- Wrote **RESTful** API for save, load and update game records using **SpringMVC** to interact with **SQLite** database.

Website Statistics Analysis using Hadoop 10/2019-01/2020

- Implemented **Hadoop Distributed File System** on Linux and use it to upload and download big scale of data (website log statistics).
- Wrote **MapReduce** functions using Java to do word count (the distribution of the total visit time per hour in one day, the number of people visit each web page, number of users in different regions) from the user log statistics.
- Designed and implemented **RESTful APIs** for backend to create, retrieve, and update the analysis result using **Spring Boot, MySQL**.
- Built the real-time data dashboard to visualize the website log statistics analysis result using eCharts..

PUBLICATION

Research about Cyber Security 11/2018-03/2020

Cyber Manhunt: Evaluation of Technologies and Practices for Effective Community Development and Maintenance

DOI link: https://doi.org/10.1007/978-3-030-39445-5_64

- Participated and presented as Author in Future of Information and Communication Conference (FICC 2020) in San Francisco.