YUTAO ZHOU

https://yutao-zhou.github.io/CV/ • yz4359@columbia.edu • https://www.linkedin.com/in/yutao-zhou/ • (805) 637-1617

EDUCATION

Columbia University New York City, NY
M.S. in Electrical Engineering Expected Dec 2023

University of California - Santa Barbara Santa Barbara, CA
B.S. in Physics Dec 2021
GPA: 3.67/4.00 Dean's Honors List in 2021 Winter.

LANGUAGE AND IT SKILLS

- Languages: Python (Proficiently operated in all of my internship and personal project).
- Computer Skills: HTML, MySQL, Shell script, MATLAB, Latex, Auto Desk CAD, EasyEDA, MS Office series.

PERSONAL PROJECTS

Used Car Data Visualization WebApplication Jun 2022 - Jul 2022
Independent Projects

- Build with streamlit, dealing large data set(365K data points) with Desk, Pandas, and NumPy for data filtering and cache data.
- Visualize data with scatter plot on heat map (with more than 100 selectable base maps), pie chart, scatter plot with trend line, with packages e.g. plotly, leafmap, pydeck.
- Added VIN lookup function with Get from NHTSA (National Highway Traffic Safety Administration)'s API.
- Designed AI key phrase extraction from listing description with spacy, and visualization with wordcloud with VIN query results (VIN query, key phrase generation, and word cloud should take less than 5 seconds, usually 2 seconds).
- Implemented geocoding and filtering data with user input distance from user query location with geoencoder in GeoPy (Entire query should take 3 seconds depends on setting usually less than 0.5 second).
- Added Login page with cookie. Hosting web application on a personal server with domain re-direction (2 month so far without offline or interference).
- Tested google map API with REQUEST for geocoding but abandoned for cost ($5 per 1000 query).
- deployed website to google cloud platform with docker and yaml files but abandoned it for daily limit of outgoing internet traffic for app engine with free account.
- Link to web application: www.auto-showroom.com login with (username: guest password: guest).

NLP Model's Comparison in Key Phrase Extraction Web Application Jun 2022 - Jul 2022
Cooperative Project

- Highlight outcome of different supervised, unsupervised and bert-unsupervised ML model's key phrases output in business report domain. Adjustable setting with customized natural language input. Total runtime under 1 second in general.
- Link to web application: http://128.111.106.229:8502.

WORK EXPERIENCE

Deepchem Co., Ltd Beijing, China
Python Intern Feb 2022 - May 2022

- Designed and builded calculation task distribution systems. Distributing calculation jobs from distribution server to different calculation servers (Group project, 4 people in total (include one manager)).
- Communicated and collaborated with front-end, and other co-walkers to create web-based platform. Represented team to communicate with manage Finished building in 1 month.
- Checked job status on platform and handled manual stop from user with GET. Handling exceptional cases e.g. distribution server offline. Stress tested on all 4 calculation servers.
- Checked front-end job status and submitted log content from calculation to distribution server in real-time with GET and POST. Zip needed calculation results and uploaded files to distribution server with POST (up to 10 jobs in real life).
- Increased overall calculation efficiency by 50% - 200% (Our platform make it possible to keep calculation servers busy on weekend and at night).
- Created algorithms finds missing data set in database from id queries CSV. Data filtering and aligning. Extract 3D Cartesian coordinate and get SMILE with Pybel(OpenBabel) python package.
- Constructed and maintain SQL database. Extract data from XYZ file, CSV file, and converted SMILE and insert it into SQL database(including checking repeating data in database)(Individual project).
- Developed an algorithm to audit two-way connections between PC and lab equipment automatically with SOCKET (Individual project finished in 1 day).