HAIZHOU LI

https://github.com/Hydr0li hydr0@bu.edu
Boston, MA 02215 (540)-998-6676 https://hydroli.top

EDUCATION

Boston University *2021.8 – Present*
MS in Computer Science In progress
Coursework: Web Analytics and Mining, Machine Learning, Data Structure, Analysis of Algorithms

Virginia Polytechnic Institute and State University *2017.8 – 2020.12*
BA in Economics (EACS) Major GPA: 3.5/4.0
Coursework: Econometrics, Statistics and Probabilities, Data Visualization and Data Cleaning with Python, Machine Learning in Economics

HONOR

Being named to the College of Science 2020 Fall Dean's List with an overloaded semester.

TECHNICAL SKILLS

Programming:	Python(sklearn, pandas, numpy, BeautifulSoup4, flair), Java, R, SQL, HTML, CSS, JavaScript, STATA, Matlab.
Machine Learning & Deep Learning:	Principle Component Analysis(PCA), Regularization, Clustering K-means, K Nearest Neighbors, Penalized regression(LASSO, Ridge) Neural Networks, Gradient Descent Algorithm, Text Mining, Embedding Documents, LDA modeling, NLP modelings, Word2vec

WORK/INTER EXPERIENCE

China Tower Xi'an Branch (Available in GitHub) *2021.2 – 2021.7*
Software Develop Engineer, Data Analytic group leader
- Normal and Logic regressions and TF-IDF matrix on Python checking whether the electricity cost reasonableness on each station.
- Helped operation and maintenance departments to formulate a strategy on local tower station.
- Outputting an .exe executable file to audit the past 3 years rationality of the electricity cost by using Python, the amount of yearly electricity cost of Xi'an City in 2021 is over 2 billion CNY. All the code available on GitHub.

China Mobile (Shaanxi) Cloud Computing Center *Summer 2019*
Data Science and After-sale Technical Problem Fix group attendee
- Assisted the marketing, sales, and after-sale technical support of cloud systems and DDoS defense systems.
- Assisted in the writing after-sale technical support workflow planning, documented workflow manuals Participated in the technical support team in the establishment project of the cloud system setup for the Bureau of Environmental Protection of Xianyang City, Shaanxi Province.

RESEARCH PROJECTS (ALL AVAILABLE IN GITHUB)

Research Assistant: Deep Learning NLP algorithm (Natural Language Processing). *2021.11 - 2022.7*
Assistant for Doctor Chen Jing (Boston University) *Mentor: Dokyun (DK) Lee (Boston University)*
- Deep Learning NLP algorithm, by using Flair, to do the embedding and predictions.
- Trained over 100 LDA Topic Embedding models to Analyze the topics from article news.
- Matching the LDA result to the tweets database by using Fuzzy and BERTopic.

Grocery Shopping App: Software development. *2022.5 – 2022.9*
Project Manager Back-End Engineer
- Cooperated with 3 team members in developing the Grocery Shopping App which is designed for enable users to determine which grocery store in their area will save them the most money.
- Scheduling meetings every period(normally twice a week), collect requirements and talk with front and back end

developers.
- As back-end engineer, successfully scraped a large scale of data from different grocery stores by using Python and built our own database by using MongoDB. This database becomes the primary data source for the application.
- Cooperated with Front-End Engineer to develop routes by using JavaScript to enable the buttons the algorithms such as calculation and check out.

Machine Learning project: Economic regression analysis on How the properties of applicants affect credit card application result. *2020.8 – 2020.12*

Individual Researcher *Mentor: Ali Habibnia (Virginia Tech)*

- With Python, scraped data from Kaggle about the credit card applications data of What factors affect the application result, and do the data cleaning and visualization.
- Designed penalized and logic regressions and PCA algorithms to train the models and conducted the analysis.

Last name categorization in census analysis *2020.8 - 2020.12*

Individual Researcher *Mentor: Melinda C. Miller (Virginia Tech)*

- By using Python to Categorized and verified last names of Indian indigenous people from the 1900 US census according to the names' language family origin.
- Designed algorithms and trained models with sklearn that involves about 833,000,000 calculations and successfully conducted the categorization and verification the last names.