Pengyi Ye

DATA ENGINEERING · SOFTWARE ENGINEERING · MACHINE LEARNING

PYE3@WPI.EDU | (508) 873-8115 | github.com/HaydenInEdinburgh

EDUCATION

Worcester Polytechnic Institute (WPI)   Worcester, MA
M.S. in Data Science   GPA:3.81/4.0   *Sep. 2018 –May. 2020*
Zhejiang University of Finance and Economics (ZUFE)   Hangzhou, China
B.S. in E-Business   GPA: 3.75/4.0   *Sep. 2014 – Jun. 2018*

SKILLS

Programming Language:	Python, Java, MySQL, Linux, JavaScript, HTML5, CSS
Technical Frameworks:	Django, Flask, MongoDB, Git, AWS
Research Domains:	ML, Natural Language Processing, Social Web

WORK EXPERIENCE

Doble Engineering Company (Graduate Qualified Project)   Marlborough, MA
Data Science Co-op   *Jan. 2020 – Present*
· Took Multiple threat and vulnerability data sources and vendor websites, to identify product-specific items.
· Extracted related records by fuzzy searching and correlated with Doble's Database to assess risk profiles.
· Developed product portfolio-specific analytic displays using **Django** and **SQLite**.

Worcester Polytechnic Institute (INFO LAB)   Worcester, MA
Research Assistant   Advisor: Kyumin Lee   *May. 2019 – Dec. 2019*
· Developed frameworks with the **Attentional LSTM** Model and **Bi-LSTM** Model on 4 million pieces of Tweets
· Multiprocessed experiments and automated the data pipeline on GPU Cloud Server with **Linux**.
· Outperform state of art **Word2Vec** embedding model i.e. GloVe and $BERT_{Base}$ with final accuracy as 94.8%.

Zhejiang DAHUA Technology Co. Ltd   Hangzhou, China
Data Engineer Intern   *Jun. 2017 – Sept. 2017*
· Quantified the problem of traffic target recognizing by analyzing graph data using python on **AWS**; Cleaned and manipulated data to create daily scores; Applying data augmentation.
· Extracted and manipulated data for recognizing the auto's label using **SQL**.
· Conducted statistical analysis via visualization in **Tableau** Interactive dashboard on model tests' performance.

Hangseng Bank Limited, Zhejiang Branch   Hangzhou, China
Finance Intern   *Mar. 2018 – July. 2018*
· Retrieved Volatility Information of the investment products from over 5,000 records to **MySQL** database.

SELECTED PUBLICATION

P. Ye, G. Mou, K. Lee, *"SWE2: SubWord Enriched and Significant Word Emphasized Framework for Hate Speech Detection"* Accepted by ***CIKM 2020****.*   *Oct. 2019*

SELECTED PROJECTS

Bernie Library   *Feb. 2020*
· Project URL: https://bernielib.herokuapp.com/
· Created a **Django** skeleton website and application, defined URL mappings, views with customized templates.
· Utilized **GitHub**'s API to post, developed Introducing Git Pages and set on production environment, **Heroku**.

Automatic Rap Maker   *Mar. 2019*
· Project URL: evakli11.github.io/CS525/
· Created a python-based program to customize and produce rap from crawled lyrics data by **LSTM** Model.
· Designed a **Tkinter** based Front-end GUI Interface to simplify using processes.