Dawei Yin

• davidyin@bu.edu • 857-316-5326 • www.linkedin.com/in/davyin/

EDUCATION

BOSTON UNIVERSITY Boston, MA

MS in **Computer Science**, concentration in **Data Analytics** Expected May 2023

• Awards/Honors: Graduate Research Assistantship & Teaching Assistantship.

• GPA: 3.83

BA in **Computer Science & Mathematics** May 2021

• Awards/Honors: *Dean's List* 3 times *(*spring of 2017, 2019, 2020).

PROFESSIONAL EXPERIENCE

HI-Lab, Boston University Boston, MA

Research Assistant Jan 2022 - present

• Implemented 5 clinical study group allocation algorithms, including urn minimization and block randomization.

• Designed and developed a management system for participant allocation utilizing Flask and MongoDB to provide group allocation service for clinical research study.

• Streamlined the participant enrollment process and allowed frictionless integration with other electronic system.

Security Department, Tencent America Palo Alto, CA

Cybersecurity Intern June 2022 – Aug 2022

• Implemented security & compliance tools utilizing Wagtail CMS, provided the compliance team a simplified process for screening security risks.

• Contributed to company internal knowledge-based system and assisted other employees with security processes such as email migration.

Technology Department, PING AN BANK CO., LTD. Shenzhen, China

Software Engineering Intern June 2020 – Sep 2020

• Created a web-crawling application utilizing scrapy to automatically migrate company internal records to long-term storage.

• Constructed a development tool to automatically integrate the crawled data into Django CMS.

PROJECTS

Real Estate Data mining Project **Boston University**

Team Lead & Researcher Sep 2021 - Dec 2021

• Led a team of 3 to research the relationship between property values, physical features and land value in Boston in the recent 10-year.

• Assessed 6 popular Data Mining models in R and Weka, including Random Forest and Decision Tree, for investigating the leading factors behind such relationship.

• Built the final model and achieved a 10% accuracy improvement on relative property value prediction, utilizing feature selection, data preprocessing, and adaptive boosting method.

Kaggle Machine Learning Competition Nov 2020 - Dec 2020

Individual Participant

• Participated in the Kaggle competition for supervised and semi-supervised plant image classification, and placed 8th out of 70 participants.

• Designed a deep learning model by utilizing transfer learning of VGG16 in TensorFlow and achieved more than 80% accuracy.

• Utilized BU Shared-Computing-Cluster (SCC) to reduce model training time.

Server Queuing Simulator Sep 2020 - Nov 2020

Researcher & Developer

• Constructed a discrete event simulator in Java to mimic various queuing models, including M/M/N and M/M/1/k, as well as different scheduling schemes, such as FCFS, SJF, and RR.

• Provided a method for analyzing the models with non-statistically significant distribution event time and time-dependent service time.

Population Migration Modeling Project May 2019 - Jun 2019

Researcher & Developer

• Led a team of 4 to investigate the population movement between city and suburb from 1995 to 2000 in China.

• Modeled the population Dynamics in more than 10 Chinese provinces utilizing MATLAB with matrix population modeling technique.

SKILLS

Languages: Python, R, Java, JavaScript, SQL, C++, Haskell, HTML/CSS, MATLAB

Tools: Git, Postman, MongoDB, GCP, Weka, Spark RDD, Hadoop, Jira

Framework: Django, Flask, Scrapy, Wagtail, TensorFlow