SHARA DUONG

(832)946-0070 – skduong@outlook.com – github.com/skduong – Houston, TX

EDUCATION

Master of Science in Statistics and Data Science — Dec 2021
University of Houston – Houston, Texas
GPA: 3.927

Bachelor of Science in Computer Science| Bachelor of Science in Mathematics — May 2019
University of Houston – Houston, Texas
Minor in Economics
GPA: 3.938

RELEVANT COURSE WORK

Algorithms and Data Structures	Big Data Analytics	Database Systems
Information Visualization	Programing for Data Analytics	Statistical Learning and Data Mining

ACADEMIC PROJECTS

Consumer Purchase Prediction, University of Houston — July 2021

- Performed predictive analytics on eCommerce data to forecast consumer purchase behaviors
- Constructed, trained, and tuned machine learning models using Spark's MLlib pipelines
- Handled computations on a large dataset using AWS S3 cloud storage and EMR big data platform

Text Image Classification, University of Houston — Apr. 2021

- Implemented a convolution neural network (CNN) to classify images of different fonts
- Reformatted image matrices into vectors using Python with NumPy and Pandas libraries
- Constructed, trained, and optimized the CNN using ML libraries such as Keras and TensorFlow
- Utilized drop-out learning to reduce model overfit

Forest Degradation Analysis, University of Houston — Mar. 2021

- Explored expected forest loss and identified potential causes using data from the US Forest Service
- Utilized Python's data science libraries for the extraction and transformation of relevant data
- Communicated findings through interactive visualizations, presenting a story at both the state and local levels with Tableau dashboards

Academic Performance Research, University of Houston — Nov. 2020

- Collaborated in a group of five and analyzed Texas Education Agency's statewide accountability report to find key determinants of student achievement
- Contributed as the main programmer, providing code and troubleshooting assistance to the team
- Collected, cleansed, and transformed the data into a usable spreadsheet
- Produced regression models in R showing the strong negative relationship between a school's proportion of socioeconomically disadvantaged students and academic performance

SKILLS & INTERESTS

Programing Languages: Python, R, C++
Tools: AWS EMR, Linux Bash, GitHub, Excel, OpenGL, SVN, Tableau, Unity
Certifications: Tableau Desktop Specialist
Interests: gaming, origami, drawing, tea, entomology, learning languages, puzzles